INVESTIGACION Y CIENCIA

Edición en español de **SCIENTIFIC AMERICAN**



LA MEMORIA DE LAS AVES

Mayo 1983
350 PTAS.

Los espacios en gris
corresponden a publicidad
en la edición impresa

Número 80
Mayo 1983

INVESTIGACION Y
CIENCIA
Edición en español de SCIENTIFIC AMERICAN

Colaboradores de este número:

Asesoramiento y traducción:

Enrique Macpherson: *La ley del Mar*; Luis Bou: *Microprogramación*; José L. Melena: *La escritura de la civilización del Valle del Indo*; Santiago Torres: *ADN mitocondrial*; Luis Mas: *El futuro del universo*; Joandomènec Ros: *La memoria de las aves que acaparan alimento*; F. Martínez González-Tablas: *Reacciones químicas oscilantes*; Luis Bou: *Temas metamágicos*; J. Vilardell: *Taller y laboratorio*.

Ciencia y sociedad

J. Modolell

Libros

Antonio Domínguez, José F. Prida, Juan Armengol y Jorge Sabater



LA PORTADA

La ilustración de la portada muestra un experimento diseñado para probar la memoria de un carbonero palustre, un pequeño pájaro europeo de la familia Páridos (véase "La memoria en las aves que acumulan alimento", por Sara J. Shettleworth, en este mismo número). Al pájaro se le coloca un parche hemisférico que le cubre un ojo; luego se le suministran pipas de girasol, que acumula escondiéndolas en el musgo que llena un gran recipiente. Más tarde se le permite que busque las semillas. Si el parche no se ha cambiado de sitio, recordará dónde se hallan escondidas las pepitas. Pero si se coloca sobre el ojo del lado opuesto de la cabeza (es decir, el que vio los lugares de almacenamiento), el resultado es totalmente distinto: no acierta a recordar. En el carbonero palustre, cada ojo envía información únicamente a un lado del cerebro. Parece, pues, que su memoria visual no puede transferirse de un hemisferio cerebral al otro. El experimento fue realizado por David F. Sherry, John R. Krebs y Richard J. Cowie.

Suscripciones:
Prensa Científica, S. A.
Calabria, 235-239
Barcelona-29 (España)
Teléfono 322 05 51 ext. 37

Condiciones de suscripción:
España:
Un año (12 números): 3.850 pesetas
Extranjero:
Un año (12 números): 52 U.S. $
Ejemplar atrasado ordinario:
350 pesetas
Ejemplar atrasado extraordinario:
450 pesetas

Distribución para España
Distribuciones de Enlace, S. A.
Ausias March, 49 - Barcelona-10

Distribución para los restantes países:
Editorial Labor, S. A.
Calabria, 235-239 - Barcelona-29

Publicidad:
Madrid:
Gustavo Martínez Ovin
Avda. de Moratalaz, 137 - Madrid-30
Tel. 430 84 81
Cataluña:
Lourdes Padrós
Manuel Girona, 61, 2.° - Barcelona-34
Teléfono 204 45 83

Controlado por O.J.D.

PROCEDENCIA DE LAS ILUSTRACIONES

Pintura de la portada de Enid Kotsehnig

Página	Fuente
11	UN Photo
12-13	Andrew Tomko
14	Walken Graphics
16	Fred N. Spiess y Peter E. Lonsdale, Institución Scripps
17	Andrew Tomko
21	Hewlett-Packard Company (*arriba*); Ilil Arbel (*abajo*)
22	Zilog, Inc.
23	Internal Revenue Service, U.S. Department of the Treasury (*arriba*); Ilil Arbel (*abajo*)
24-28	Ilil Arbel
31	James P. Blair, cortesía del Museo Nacional de Pakistán
32	James P. Blair, cortesía del Museo Nacional de la India (*arriba*); Museo Americano de Historia Natural (*abajo*)
33-37	Andrew Christie
38	McGuire Gibson, Oriental Institute, Universidad de Chicago
39	Andrew Christie
47	Ernst F. J. Van Bruggen, Universidad estatal de Groningen (*arriba*); Annika C. Arnberg y Ernst F. J. Van Bruggen, Universidad estatal de Groningen (*abajo*)
48-52	Bunji Tagawa
53	Annika C. Arnberg y Ernst F. J. Van Bruggen, Universidad estatal de Groningen (*arriba*); Bunji Tagawa (*abajo*)
54-56	Bunji Tagawa
58	J. H. Hoeimakers, Universidad de Amsterdam
61	Ann Savage, Royal Observatory Edinburgh, Coonabarabran, Australia
62-70	Walken Graphics
73-74	Tom Prentiss
75	Stephen B. Vander Wall, Universidad estatal de Utah
76-81	Tom Prentiss
83	R. F. Bonifield
84-90	Allen Beechel
92	R. F. Bonifield
95	F. Rodríguez Valera y F. Ruiz Berraquero
96	Ricardo Génova, F. Rodríguez Valera y F. Ruiz Berraquero (*arriba*); F. Rodríguez Valera y F. Ruiz Berraquero (*abajo*)
97-102	Ricardo Génova, F. Rodríguez Valera y F. Ruiz Berraquero
105-109	Edward Bell
110	Ilil Arbel
113-118	Michael Goodman

ISSN 0210-136X
Dep. legal: B. 38.999-76
Fotocomposición Tecfa
Pedro IV, 160 - Barcelona-5
Fotocromos reproducidos por GINSA, S.A.
Imprime GRAFESA
Gráfica Elzeviriana, S. A.
Napoles, 249 - Tel. 207 40 11
Barcelona-13
Printed in Spain - Impreso en España

Los autores

ELISABETH MANN BORGESE ("La Ley del Mar") es profesora de ciencias políticas de la Universidad de Dalhousie en Nueva Escocia (Canadá). Hija del novelista Thomas Mann, nació en Munich y se licenció por el Conservatorio de Música de Zurich en 1937. En 1938 emigró a los Estados Unidos con su familia, adoptando la nacionalidad norteamericana en 1941. Antes de ingresar en la Universidad de Dalhousie, en 1978, trabajó en el Centro de Estudios de Instituciones Democráticas de Santa Bárbara. Se ha especializado en derecho marítimo, del que ha escrito numerosos artículos y libros.

DAVID A. PATTERSON ("Microprogramación") enseña ingeniería eléctrica y ciencias de cómputo en la Universidad de California en Berkeley. Antes de doctorarse, en 1976, trabajó para la Hughes Aircraft Company en el diseño y evaluación de ordenadores. Ha permanecido en Berkeley desde 1977. En el otoño de 1979 colaboró con la Digital Equipment Corporation, en la puesta a punto de técnicas para diseñar microprogramas. En uno de sus cursos de Berkeley proyectó, con sus alumnos, un microprocesador llamado RISC I (Reduced-Instruction-Set Computer I), donde todos los programas se convierten en microprogramas.

WALTER A. FAIRSERVIS, JR. ("La escritura de la civilización del Valle del Indo") da clases de antropología en el Vassar College. Cursó el primer ciclo de los estudios de arte en la Universidad de Chicago. Se licenció por la de Columbia y se doctoró en antropología por la de Harvard. Su carrera profesional ha compaginado la docencia con la dirección de campañas antropológicas en Medio y Lejano Oriente. Antes de ingresar en Vassar, en 1968, enseñó en las universidades de Nueva York y Washington. Fairservis escribe: "Tengo la creencia implícita de que, en los orígenes de la civilización, nos la estamos viendo con algo que no acabamos de comprender".

LESLIE A. GRIVELL ("ADN mitocondrial") nació y se educó en Inglaterra. Licenciado en ciencias en 1966, se doctoró en bioquímica, en 1969, por el University College de Londres. Nos cuenta: "Me interesé por la biogénesis mitocondrial en mis años estudiantiles y tuve la fortuna de entrar en ese campo en su momento de esplendor, a finales de la década de 1960 y principios de la de 1970". Desde 1969 trabaja en el laboratorio de bioquímica de la Universidad de Amsterdam, donde actualmente enseña biología macromolecular.

DUANE A. DICUS, JOHN R. LETAW, DORIS C. TEPLITZ y VIGDOR L. TEPLITZ ("El futuro del universo") son físicos interesados en las relaciones entre la física de altas energías y la cosmología. Dicus enseña física en la Universidad de Texas en Austin. Cursó estudios universitarios en Washington y se doctoró por la Universidad de California en Los Angeles (UCLA). Investigó en UCLA, en el Instituto de Tecnología de Massachusetts (MIT) y en la Universidad de Rochester, trasladándose luego a la de Texas, en 1973. Letaw es director técnico de la Sociedad de Comunidades Severn, en Severna Park, Maryland. Licenciado por la Universidad de Clark, en 1975, se doctoró en física por la Universidad de Texas en Austin. Doris Teplitz es profesora de la Universidad de Maryland en College Park. Comenzó su formación en el Wellesley College, aunque recibió el doctorado en física por la Universidad del Nordeste. De 1974 a 1978 formó parte del claustro del Instituto Politécnico y de la Universidad estatal de Virginia. Vigdor Teplitz es jefe en funciones de la división de asuntos estratégicos de la Oficina de Control de Armas y Desarme de los Estados Unidos y profesor adjunto de física de la Universidad de Maryland. Cursó la licenciatura en el MIT y, el doctorado, en la Universidad de Maryland. Trabajó en el MIT y en el Instituto Politécnico y Universidad estatal de Virginia. En 1979 fue consejero de la delegación de los Estados Unidos en las sesiones segunda y tercera de las negociaciones sobre proyectiles contrasatélites mantenidas entre los Estados Unidos y la Unión Soviética.

SARA J. SHETTLEWORTH ("La memoria de las aves que acaparan alimento") enseña psicología en la Universidad de Toronto. Cursó el primer ciclo de estudios universitarios en el Swarthmore College; prosiguió su formación en la Universidad de Pennsylvania y se doctoró en psicología, en 1970, por la de Toronto. Cuenta la autora: "Mi tesis doctoral describe los distintos efectos que ejerce el premio y el castigo en actividades como el acicalamiento, la excavación y el encabritamiento de los hámster dorados. Ultimamente me interesa también el papel que desempeñan los mecanismos de aprendizaje y memoria en la búsqueda de alimento. He trabajado igualmente en el estudio de la conducta de las tortugas de mar, con Nicholas Mrosovsky, mi marido".

IRVING R. EPSTEIN, KENNETH KUSTIN, PATRICK DE KEPPER y MIKLOS ORBAN ("Reacciones químicas oscilantes") son químicos que comparten una misma inquietud: el tema sobre el que versa su artículo. Epstein es profesor de la Universidad de Brandeis. Procede de Harvard. Ingresó en Brandeis en 1971. En 1977 y 1978 visitó el Instituto Max Planck de Biofísica química, de Göttingen, becado por la Fundación Nacional de la Ciencia. Kustin enseña también en Brandeis. Se licenció por el Queens College, en 1955, y se doctoró en química inorgánica por la Universidad de Minnesota. Becado por el Servicio de Salud Pública de los Estados Unidos estuvo en el Instituto Max Planck de Química-Física de Berlín, desde donde se trasladó, en 1961, a Brandeis. De 1974 a 1977 fue secretario del departamento de química. De Kepper investiga sobre ingeniería química en el centro Paul Pascal de la Universidad Domaine, adscrito al Centro Nacional francés de Investigaciones Científicas (CNRS). Nacido en Francia, se educó en la Universidad de Burdeos I, recibiéndose de doctor en 1978; ese mismo año se incorporó al laboratorio del CNRS. Orbán nació y se educó en Hungría. Doctor por la Universidad de Eötvös, forma parte de su claustro docente desde 1962.

FRANCISCO RODRIGUEZ VALERA y FRANCISCO RUIZ BERRAQUERO ("Halobacterias") se hallan adscritos al departamento de microbiología de la Facultad de Farmacia de la Universidad de Sevilla, donde el segundo desempeña el cargo de catedrático y, de profesor agregado, el primero. Rodríguez Valera se doctoró en biología por la Universidad de Granada. Ha investigado en el Laboratorio de Biología Marina de Woods Hole y en el departamento de biología de la Universidad de Ottawa. Ha descubierto la vía de dispersión de los halófilos extremos. Ruiz Berraquero se doctoró en biología por la Universidad Complutense en 1967, aunque inició sus estudios en la de Granada. Dirige las investigaciones sobre microorganismos halófilos que se están llevando a cabo en su departamento hispalense.

Hace…

José M.ª López Piñero

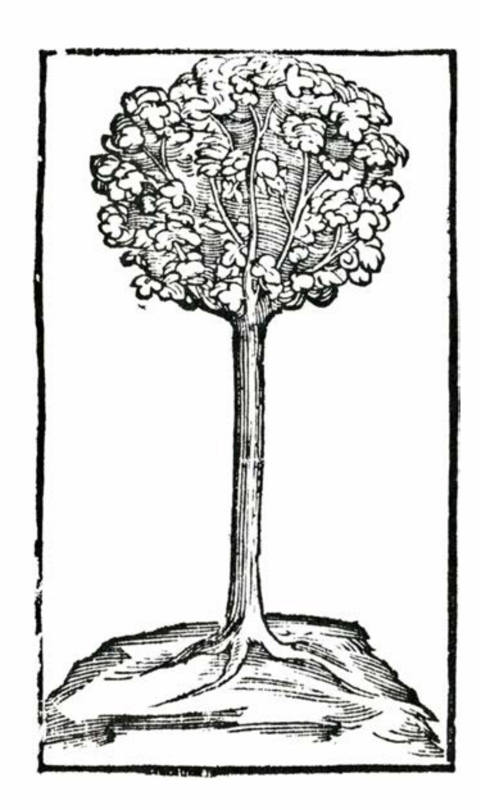
… cuatrocientos cincuenta años

Nicolás Bautista Monardes obtuvo el grado de bachiller en medicina por la Universidad de Alcalá. Allí recibió la influencia de la corriente encabezada por Elio Antonio de Nebrija, aunque el gran humanista no fue profesor suyo. En 1547 se doctoró en la Universidad de Sevilla, ciudad en la que permaneció hasta su muerte, a muy avanzada edad. Además de ejercer la medicina con gran prestigio y notables ingresos, participó en empresas mercantiles, entre ellas el comercio de materias medicinales y el tráfico de esclavos.

Monardes publicó numerosas obras. La primera de ellas, *Diálogo llamado Pharmacodilosis* (1536), refleja, como ha aclarado Pérez Fuenzalida, su adscripción a la corriente humanística. Atribuye a los árabes la decadencia de la botánica y de la materia médica y recomienda el estudio directo de los clásicos, especialmente de Dioscórides, en la misma línea de Nebrija. Parecida orientación tiene su segundo libro (1539), que trata de la sangría en el "mal de costado", tema entonces de una agria polémica entre los seguidores de la tendencia humanística y la arabizante. En fechas posteriores, publicó un estudio sobre las rosas y los frutos cítricos (*ca*. 1540), una traducción de la *Sevillana Medicina* (1545) del judío bajomedieval Juan de Aviñón, una monografía sobre la nieve (1571) y el *Diálogo del Hierro* (1574) al que después nos referiremos. Su libro más impor-

Parte ſegunda. Delas coſas que traen

EL SASSAFRAS.

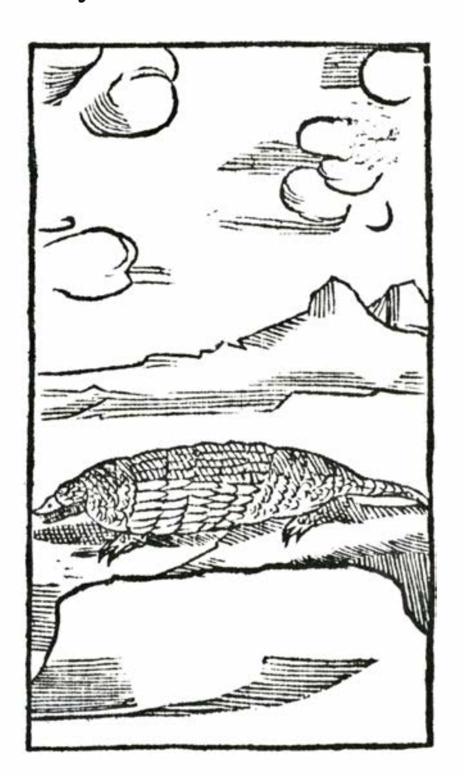


DEL

1. Página de la Historia Medicinal, *de Monardes, en la que aparece el primer grabado conocido sobre el sasafrás*

de nueſtras Indias: que ſiruen al vſo de Medicina. 66

EL ARMADILLO.



ESTE Animal ſaque de otro natural, que eſta enel Muſeo de Gonçalo de Molina: vn cauallero deſta ciudad, enel qual ay mucha cãtidad de libros de varia lectiõ, y muchos generos de animales y aues, y otras coſas curioſas, traydas aſsi dela India Oriental, como Occidẽtal, y de otras partes del mundo, y gran copia de monedas y piedras antiguas, y diferencias de armas, que con gran curioſidad y con generoſo animo ha allegado.

K 2 Traen

2. Otra página del tratado de Monardes, ésta dedicada al armadillo que tenía Argote de Molina en su museo

tante fue, sin embargo, el que en su primera edición completa se tituló *Historia Medicinal de las cosas que se traen de nuestras Indias Occidentales* (1574). Consta de tres partes, la primera de las cuales había aparecido en 1565 y la segunda en 1571.

Monardes tuvo clara conciencia del servicio que prestaba al dedicar largos años a la preparación de esta obra. En el prefacio de su primera parte, tras enumerar la "riqueza increíble" que viene del Nuevo Mundo en oro y plata, piedras preciosas, perlas, animales, "lana, algodón, grana para teñir, cueros, azúcares, cobre, brasil y ébano", afirma que todavía más valiosas son "las cosas que traen de nuestras Indias Occidentales que sirven en medicina". Por ello, se propuso estudiarlas, "de lo cual seré el primero, para que los demás añadan con este principio lo que más supieren y por experiencia más hallaren". Como Sevilla, prosigue, "es puerto y escala de todas las Indias Occidentales... púdele hacer, juntamente con la experiencia y uso de ellas de cuarenta años que ha que curo en esta ciudad, donde me he informado de los que de aquellas partes las han traído con mucho cuidado y las he experimentado".

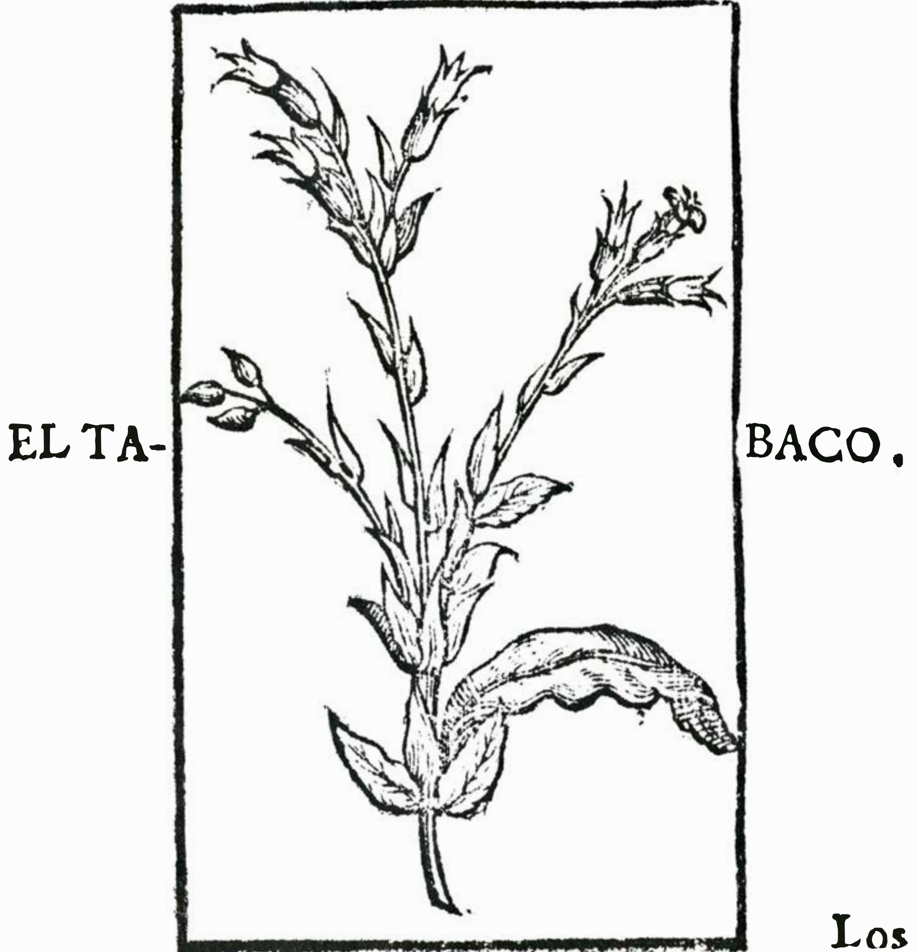
El médico sevillano no era un observador ocasional ni un autodidacta, sino

31

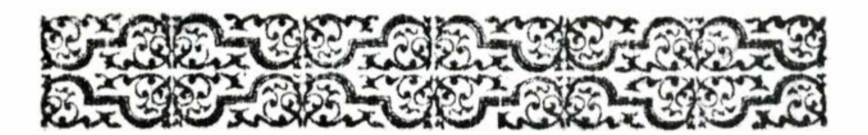
SEGVNDA PARTE

Del Libro, delas cosas que se traen de nuestras Indias Occidentales, que siruen al vso de medicina. Do se trata del Tabaco, y dela Sassafras, y del Carlo santo: y de otras muchas Yeruas, y Plantas, Simientes, y Licores: que nueuamente han venido de aquellas partes, de grandes virtudes, y marauillosos efetos.

¶ *Hecho por el Doctor Monardes: Medico de Seuilla.*



3. "El nombre propio suyo entre los indios es picietl, *que el de tabaco es postizo, de nuestros españoles, por una isla do hay mucha cantidad de él, llamada este nombre de Tabaco. Hayla y nace en muchas partes de las Indias, ordinariamente en lugares húmedos y sombríos; es menester que sea la tierra bien cultivada do se sembrare, y que sea tierra libre. Siémbrase en todo tiempo en las tierras calientes y en todo tiempo nace. En las frías se ha de sembrar por el mes de marzo, porque se defienda de las heladas. Es hierba que crece y viene a muchas grandeza, muchas veces a ser mayor que un limonero. Echa un tallo desde la raíz que sube derecho sin declinar a ninguna parte; echa muchos virgultos (= arbolitos) derechos que casi se igualan con el tallo principal. Su hoja es casi como de cidrón; salteadas, vienen (las hojas) a mucha grandeza, en especial las baxas, que son mayores que de romaza; son de color de un verde verdoso, y de este color es toda (la planta). Es vellosa la planta y sus hojas. Puestas en las paredes enjardinan como los cidrones y naranjos, porque todo el año está verde y tiene hojas; si alguna se secan son las baxas. En lo alto de toda la planta echa la flor, la cual es a modo de campanilla blanca y en medio encarnada, que tiene hasta gracia; cuando se secan parecen dormideras negras; en ella está inclusa la simiente, que es en extremo menuda, de color leonado oscuro. La raíz tiene gruesa conforme a la grandeza de la planta, dividida en gajos y es lignosa; la cual, partida, tiene el corazón de dentro a la manera de color de azafrán, y gustada tiene algún amargor. Apártase fácilmente de la corteza della."*







ELOGIO HECHO POR EL ILLVST. S. GONC,ALO C,ATI-
eco de Molina, al Retrato del Autor que
ſe vee, en ſu Muſeo.

A 3

4. Retrato de Monardes que figura en su Historia Medicinal, *encabezando un poema en su honor escrito por Gonzalo Argote de Molina*

un científico sólidamente formado y con amplia experiencia en el estudio de la naturaleza. Cultivó plantas americanas en el huerto de su casa y aprovechó las colecciones y jardines que existían entonces en Sevilla, entre ellos el museo de Gonzalo Argote de Molina y quizá también el jardín botánico de Simón de Tovar. Describió por vez primera, según Francisco Guerra, varias especies vegetales, como la amapola espinosa (*Argemone mexicana*), la cebadilla (*Sabadilla officinarum*), la jalapa (*Exogonium purga*) o el sasafrás (*Sassafras officinale*), pero, sobre todo, ofreció las primeras descripciones detalladas y correctas de otras muchas. Por ejemplo, se consideran clásicas las que dedicó al tabaco –del que publicó el primer grabado– y de los bálsamos medicinales más importantes: el llamado del Perú (*Myroxylon pereirae*) y el de Tolú (*Myroxylon toluifera*). Familiarizó también a los europeos con otras plantas de gran importancia dietética o medicinal, como el maíz, la piña tropical, el cacahuete, la batata, la zarzaparrilla o la coca. Mucho más escasas son sus referencias a especies animales, una de las cuales está dedicada al armadillo, sobre la base de un ejemplar vivo que poseía Argote de Molina. Entre las que consagra a minerales destacan las correspondientes al ámbar, al azufre procedente de Quito y Nicaragua y al origen del bitumen y del petróleo.

El aspecto más brillante de su obra es, sin embargo, su contribución a la farmacognosia. "Monardes –afirma Guerra– rebasa el antiguo concepto de los herbarios tradicionales para ahondar en el origen y caracteres de identificación además de los usos de los materiales americanos, apuntando la similitud y diferencias con los productos de la farmacia clásica." A ello se debe que, desde el ya clásico *Handbuch der Pharmakognosie* (1933), de Alexander Tschirch, se le considere, junto a Valerius Cordus y a Charles de l'Escluse (Clusius), como uno de los "padres" de esta ciencia.

El principal interés del *Diálogo del Hierro* (1574), que antes hemos citado, reside en que fue el primer libro consagrado a un tema que no volvería a ser tratado monográficamente hasta comienzos del siglo XVIII. Monardes defiende en él la importancia del hierro para la cultura humana, contraponiéndolo a la inmoderada ambición que despertaban los metales preciosos. La exposición se hace en forma de un diálogo entre el propio autor, el boticario Bernardino de Burgos –personaje real según Francisco Guerra– y un metalurgista vizcaíno llamado Ortuño. En boca de este último se pone la descripción de los principales yacimientos europeos y españoles, así como de las técnicas de labrado del hierro y de la fabricación del acero. Sin demasiada profundidad, se recogen principalmente los métodos habituales en las famosas "ferrerías" del País Vasco. El libro se ocupa igualmente de la génesis del hierro, de sus cualidades y de sus aplicaciones terapéuticas, con un enfoque propio de la cultura académica de la época.

Las obras de Monardes se encuentran entre los textos científicos del siglo XVI que más interés despertaron en la Europa de la época. En poco más de cien años, alcanzaron cuarenta y dos ediciones en seis idiomas. A esta difusión contribuyeron de modo decisivo las traducciones latinas, resumidas y comentadas, que realizó el propio Clusius.

La Ley del Mar

En diciembre de 1982, 119 naciones firmaron la Convención sobre la Ley del Mar, que proporciona un amplio marco para la protección del medio marino, la investigación oceánica y la resolución de litigios

Elisabeth Mann Borgese

El 6 de diciembre de 1982, los representantes de 119 naciones, reunidos en Montego Bay, Jamaica, firmaron la Convención de las Naciones Unidas sobre la Ley del Mar. La acción coronaba 15 años de labor y daba fin a la más concurrida y prolongada de todas las conferencias internacionales celebradas. A la Tercera Conferencia de las Naciones Unidas sobre la Ley del Mar asistieron unos 3000 delegados; se desarrolló a lo largo de 11 sesiones que totalizaron 585 días. A partir de una situación caótica y superando las luchas propias de intereses políticos y económicos encontrados, corrigiendo borrador tras borrador, la Conferencia consiguió escribir una "constitución para los océanos", que consta de 320 artículos (agrupados en 17 partes principales) y nueve anexos técnicos.

Si se nos permitiera quintaesenciar tamaño trabajo en una sola frase, diríamos que la Convención sustituye el tradicional *laissez faire* que gobernaba las relaciones marítimas por un nuevo sistema de administración. Se acabó el mar ilimitado de dos dimensiones; la Convención reparte tridimensionalmente sus fondos, de gran interés económico, así como la superficie. Concede el 40 por ciento del océano y fondos adyacentes a las costas de todos los continentes e islas a la administración de los estados ribereños. Reserva el 60 por ciento restante para el sistema tradicional de libertad en el mar, pero conservando la riqueza de los océanos, que supone el 42 por ciento de la superficie terrestre, para la llamada "herencia común de la humanidad". La Convención pone esta herencia bajo la administración de una nueva autoridad internacional de los fondos marinos, con capacidad para generar riqueza, establecer cánones de captura y una especie de dominio sobre las técnicas empleadas en la explotación de los océanos.

Aunque sin oponerse a la autoridad que rige los fondos marinos, los Estados Unidos y otras 15 poderosas naciones industrializadas (entre ellas España) negaron su firma a la Convención. Sí suscribieron el acta final de la Conferencia (solamente Turquía negó su firma a ambas). Las naciones que firmaron el acta final, pero no la Convención, podrán asistir a las reuniones futuras de las comisiones creadas por la Convención, aunque en calidad de observadores. La firma de la Convención da a ésta la fuerza de una ley, sujeta a la ratificación de, al menos, 60 naciones. Cuando reciba tal ratificación, cabe esperar que los ahora no firmantes se adhieran como miembros de pleno derecho.

La Convención proporciona un amplio marco para la protección del medio marino, un nuevo régimen para la investigación marina y un amplio sistema para la resolución de litigios. Asegura la libertad de navegación y el libre paso por los estrechos surcados en la navegación internacional, invocando un nuevo concepto legal de "tránsito", un derecho que no puede suspenderse bajo ninguna circunstancia. Actualiza y codifica las leyes tradicionales; define los derechos y los deberes de los estados y organizaciones internacionales, a fin de establecer y mejorar un nuevo orden en el mar y en los océanos.

En la elaboración de la nueva ley se han contemplado todos aquellos temas que debe afrontar la comunidad mundial de nuestro tiempo: alimentación y energía, carrera de armamentos, comunicación y "revolución de la información", comercio, administración y conservación de los recursos, integración regional y los progresos de la ciencia y la técnica en conexión con los puntos anteriores. Es como si todos esos aspectos convergieran en el océano, y las aguas fueran ahora el laboratorio donde construir un nuevo orden mundial. Ese orden, esperamos, será más racional, más humano y con mayor capacidad de respuesta ante las necesidades reales del mundo que el orden antiguo, que se desmorona entre el hambre y la violencia.

Hasta 1973, cuando dio comienzo la Tercera Conferencia, las leyes del mar habían sido siempre dictadas por uno, o dos o media docena de países costeros. A través de los siglos, o de los milenios, el mundo antiguo se rigió por el código de navegación de Hammurabi y, más tarde, por las leyes del mar de Rodas. El Mediterráneo, cuna de la civilización occidental, fue el *mare nostrum* de Roma. En 1493, España y Portugal se dividieron "la mar océana" allende el estrecho de Gibraltar. Para la potencia marítima del siglo XVII, Gran Bretaña, John Selden propugnó la teoría del *mare clausum*, mar cerrado y propiedad de los estados costeros, que ejercerían el control, como anteriormente España y Portugal, de la navegación en sus aguas. Contra esa doctrina, el jurista holandés Hugo Grotius defendió la idea de un *mare liberum*, un océano abierto y libre para los navegantes de todas las naciones que respetasen los derechos mutuos. La realidad histórica enseña que tal libertad sólo podían ponerla en práctica las naciones fuertes; en la historia contemporánea, sólo un puñado de países de Europa occidental y Norteamerica.

El vertiginoso desarrollo de dos episodios, concurrentes y a menudo conflictivos, hicieron inevitable la llamada a la Tercera Conferencia de las Naciones Unidas sobre la Ley del Mar. El primero fue la dimensión marina de la revolución científico-industrial que transformó las sociedades industriales y afectó a las relaciones de las sociedades industrializadas y preindustrializadas. El descubrimiento de la expansión de los fondos marinos ha incidido, cual terremoto, en los cimientos de la geología, instalando la tectónica de placas en

la base de esa ciencia. El descubrimiento tiene implicaciones económicas a largo plazo: donde se forma el fondo, en los ejes de expansión, afloran a la superficie metales separados selectivamente del manto terrestre. Se sabe ahora que por ese proceso se depositaron las vetas de metales no ferrosos que se explotan en las minas de rocas duras. Por el mismo procedimiento conocemos que el mar contiene más petróleo y gas, tanto en las plataformas continentales como en los fondos oceánicos, que los continentes (la mayor parte de los cuales fue, a su vez, fondo marino).

Mientras tanto, el rápido desarrollo técnico de la industria pesquera ha desmentido la idea ancestral de que el mar era inagotable. La pretendida fecundidad de los océanos ha recibido otro duro golpe: la contaminación resultante de la urbanización, la industrialización y el aumento de la población.

De esta forma, los supuestos de la antigua ley fueron perdiendo vigor, lo que aceleró su desmoronamiento. Los Estados Unidos empezaron a reclamar sus derechos oceánicos y sus recursos en 1945, con las Proclamaciones sobre la Plataforma Continental y la Extensión de la Zona Pesquera, de Truman. Una política de rapiña iba a dividir el océano, como ya ocurrió con Africa un siglo antes.

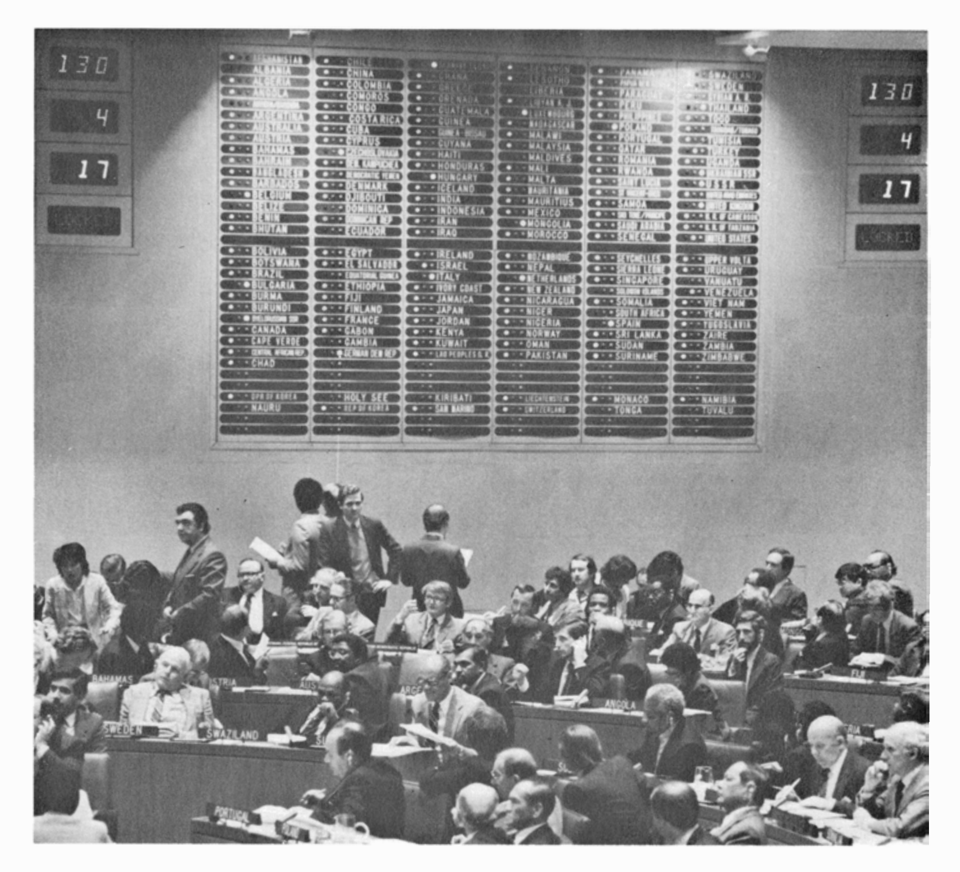
Mientras la Revolución Industrial iba penetrando en el océano por el hemisferio Norte, el segundo impulso que ha conformado la Ley del Mar cobraba fuerza en el hemisferio Sur. De los imperios coloniales de las naciones marítimas europeas comenzaron a desmembrarse nuevas naciones soberanas. No más de 70 naciones habían participado en la primera conferencia de las Naciones Unidas sobre la Ley del Mar, convocada en 1958, o en la segunda, en 1960. En 1973, cuando empezó la tercera conferencia, los miembros de las Naciones Unidas habían doblado su número. El sector mayoritario de la Asamblea General estaba ya compuesto por las naciones pobres de Asia, Africa e Iberoamérica. Su voz iba a escucharse en la Conferencia; como se oiría su llamada a un nuevo orden económico. La nueva Ley del Mar debía tenerlo en cuenta, e incluso modelarlo.

La configuración de la Convención sobre la nueva Ley del Mar muestra claramente la incidencia de las revoluciones tecnológicas y de las relaciones internacionales, de cuño paralelo, aunque a menudo enfrentadas. Fue la revolución tecnológica procedente del norte la que presionó por la expansión de la jurisdicción nacional en los océanos. Por su parte, la revolución en las relaciones internacionales instó, desde el sur, el establecimiento de nuevas y poderosas instituciones internacionales. Las dos revoluciones se plasman igualmente en las actas de tan prolongada conferencia. Las reuniones y las actas reflejan también la presencia de numerosas personalidades, muy capacitadas y eruditas, tanto del norte como del sur: Arvid Pardo, del estado insular de Malta; Hamilton Shirley Amerasinghe, de Sri Lanka, presidente de la Conferencia hasta que falleció, en 1981; Jens Evensen, de Noruega; Jorge Castañeda, de México; S. P. Jagota, de la India y Louis Sohn, de los Estados Unidos, por nombrar unos pocos. (La delegación española estuvo encabezada, sucesivamente, por los embajadores D. Antonio Poch, D. Eduardo Ibáñez y García de Velasco y D. José Manuel Lacleta.) Su buen hacer, más las presiones conjuntas del norte y del sur, guiaron los trabajos de la Conferencia.

Fue Pardo quien propuso, en una memorable intervención ante el Primer Comité de las Naciones Unidas, en 1967, que la riqueza de los fondos oceánicos debía declararse herencia común de la humanidad. Ni la libertad en el mar ni la soberanía que reclamaban los países costeros, dijo, podían ahuyentar el triple peligro de la contaminación, agotamiento de la vida marina y contiendas internacionales. Sólo la cooperación internacional podría mantener la paz y la unidad de los océanos, y ello en gran parte porque los intereses creados no estaban tan firmemente afianzados en el mar como en tierra.

Inmediatamente se creó un comité sobre la explotación del fondo marino con fines pacíficos; en 1970 se efectuó una declaración de principios que contemplaba las tesis principales de Pardo, y en 1972 se incorporó, al tratado de prohibición de empleo de armamento nuclear en el fondo oceánico, la conservación de los fondos marinos para fines pacíficos. En 1972, el comité de fondos marinos, que proseguía trabajando, apuntó la conveniencia de una Tercera Conferencia de las Naciones Unidas sobre la Ley del Mar. Se empezó a preparar el desarrollo de la misma.

Los planes terminaron en 1973, con la inauguración de los trabajos de la Conferencia. Estructurada en un comité central, tres comités principales de trabajo y un complejo número de gru-

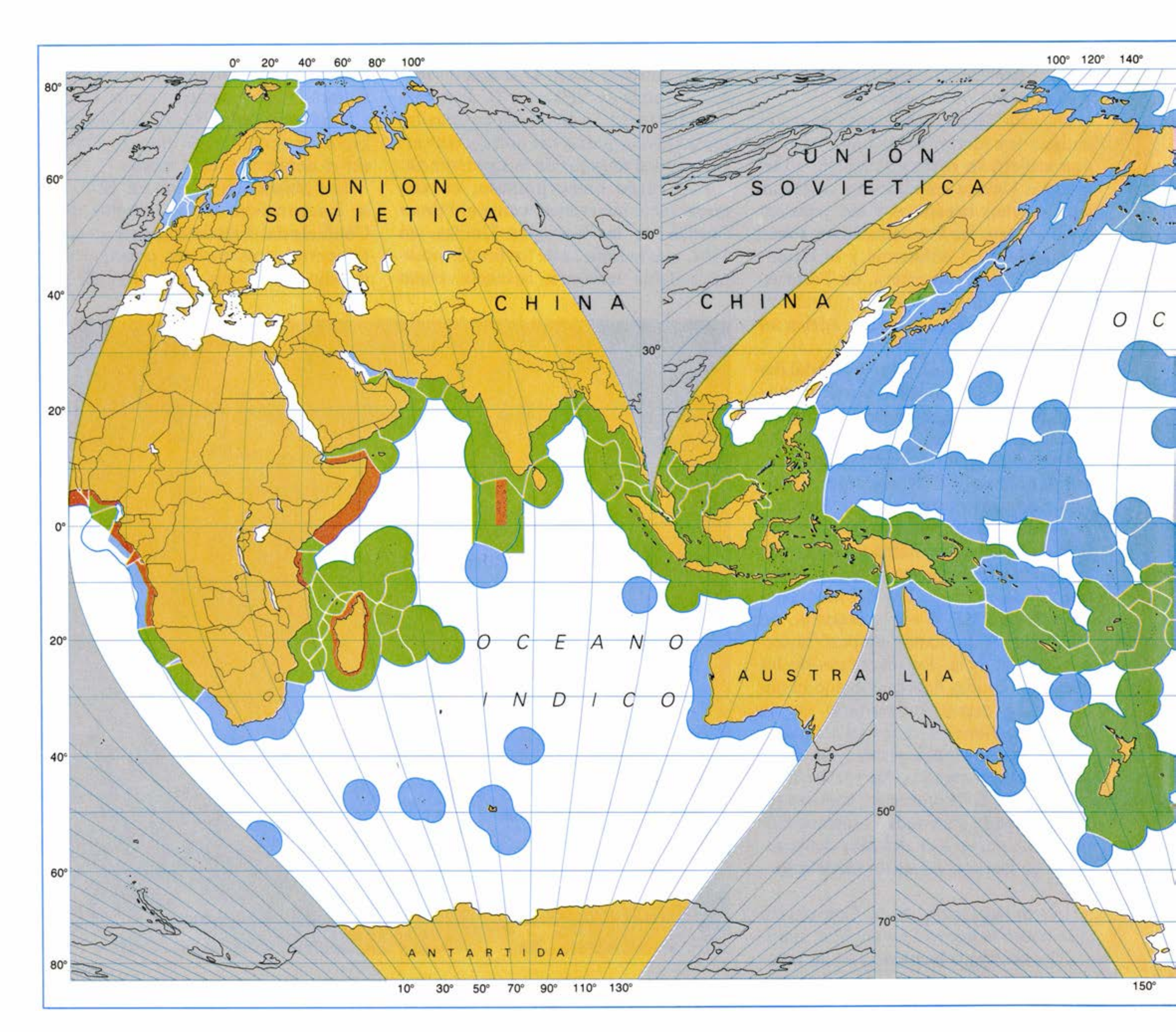


1. REUNION DE LA CONVENCION sobre la Ley del Mar, fotografiada en la Sede Central de las Naciones Unidas el 30 de abril del pasado año. Sobre los delegados, la pizarra con el recuento de los votos: 130 naciones a favor, cuatro en contra y 17 abstenciones (España se abstuvo). El 6 de diciembre, en Montego Bay, Jamaica, 119 países firmaron la Convención y 143 el acta final de la Conferencia sobre la Ley del Mar. (España no firmó la Convención pero sí el acta final.) La firma del acta final pero no de la Convención permite asistir a las futuras reuniones de la comisión en calidad de observador sin voto.

pos regionales, grupos de interés, de negociación, sesiones plenarias, sesiones plenarias informales y un comité de redacción que trabajaba en seis lenguas, ha sido la conferencia más nutrida y compleja de la historia.

La delimitación de los espacios oceánicos –las tradicionales aguas territoriales y su zona contigua, así como la nueva extensión del área económica y de la plataforma continental– correspondió al poderoso segundo comité. Debían aquí reconciliarse los intereses en conflicto de tres grupos de naciones. Los estados costeros, por supuesto, pretendían extender los límites de su jurisdicción lo más lejos posible y asegurar al máximo sus derechos en dicha zona. En el otro extremo del espectro quedaban las naciones interiores, entre ellas numerosas africanas sin salida al mar, las más pobres entre las pobres. No podían reclamar zona alguna, y se resistían a perder los derechos que hasta entonces pudieran disfrutar, particularmente aquellos referidos a la pesca, en los mares que pronto les quedarían vedados. En medio pugnaban los estados marítimos, que compartían los intereses de los estados costeros, pero que, por otra parte, dependían de la libertad de los mares para organizar sus flotas de gran altura, y por tanto se situaban, en algunos aspectos, del lado de los países interiores. Cada grupo era, además, un complejo agregado de países desarrollados y en vías de desarrollo y naciones de mercado libre y economías dirigidas.

No cabe duda de que los ganadores en esta expansión de la jurisdicción nacional han sido los países industrializados, particularmente los Estados Unidos, Canadá, Australia y Sudáfrica, así como algunos grandes países en vías de desarrollo: India, Brasil y México, entre otros. Los artículos de la Convención surgidos del segundo comité confirman el límite territorial de las 12 millas, más una zona contigua de otras 12. Más allá de esos límites se acuerda que la soberanía sobre los recursos naturales por parte de los estados costeros se



2. RECLAMACIONES MARITIMAS de diversas naciones, reflejadas en este mapa basado en el que preparó F. Webster McBryde, de Transemantics Inc. Las áreas en azul son zonas de pesca, donde se reclaman derechos pesqueros exclusivos. Las áreas en verde son zonas económicas, donde se han reclamado no sólo derechos pesqueros, sino también derechos de soberanía sobre todos los recursos y jurisdicción sobre investigación y contaminación

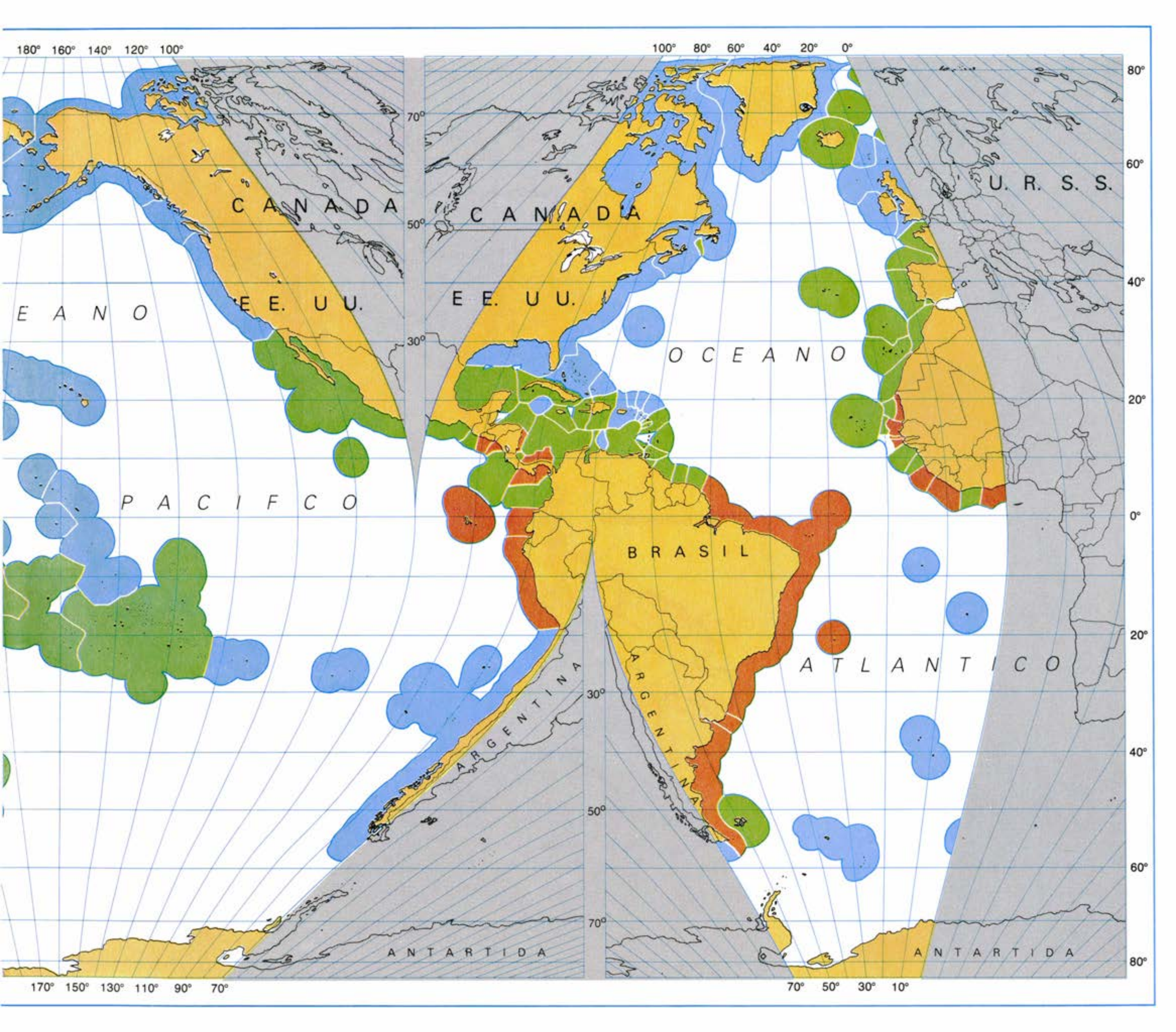
extienda hasta las 200 millas, así como una serie de derechos sobre aspectos económicos y de investigación científica en dicha zona. Donde la plataforma continental trasciende la zona económica de 200 millas, los derechos sobre los recursos de la plataforma pueden extenderse, sin sobrepasar en ningún caso las 350 millas.

Los estados ribereños no tienen soberanía sobre los territorios de la zona económica, como la poseen sobre sus tierras y territorios marinos adyacentes. Este nuevo tipo de soberanía funcional debe considerarse la forma menos mala de expansión nacional. Es un concepto nuevo y dinámico de la legislación internacional, que habrá de hallar otras aplicaciones en el futuro.

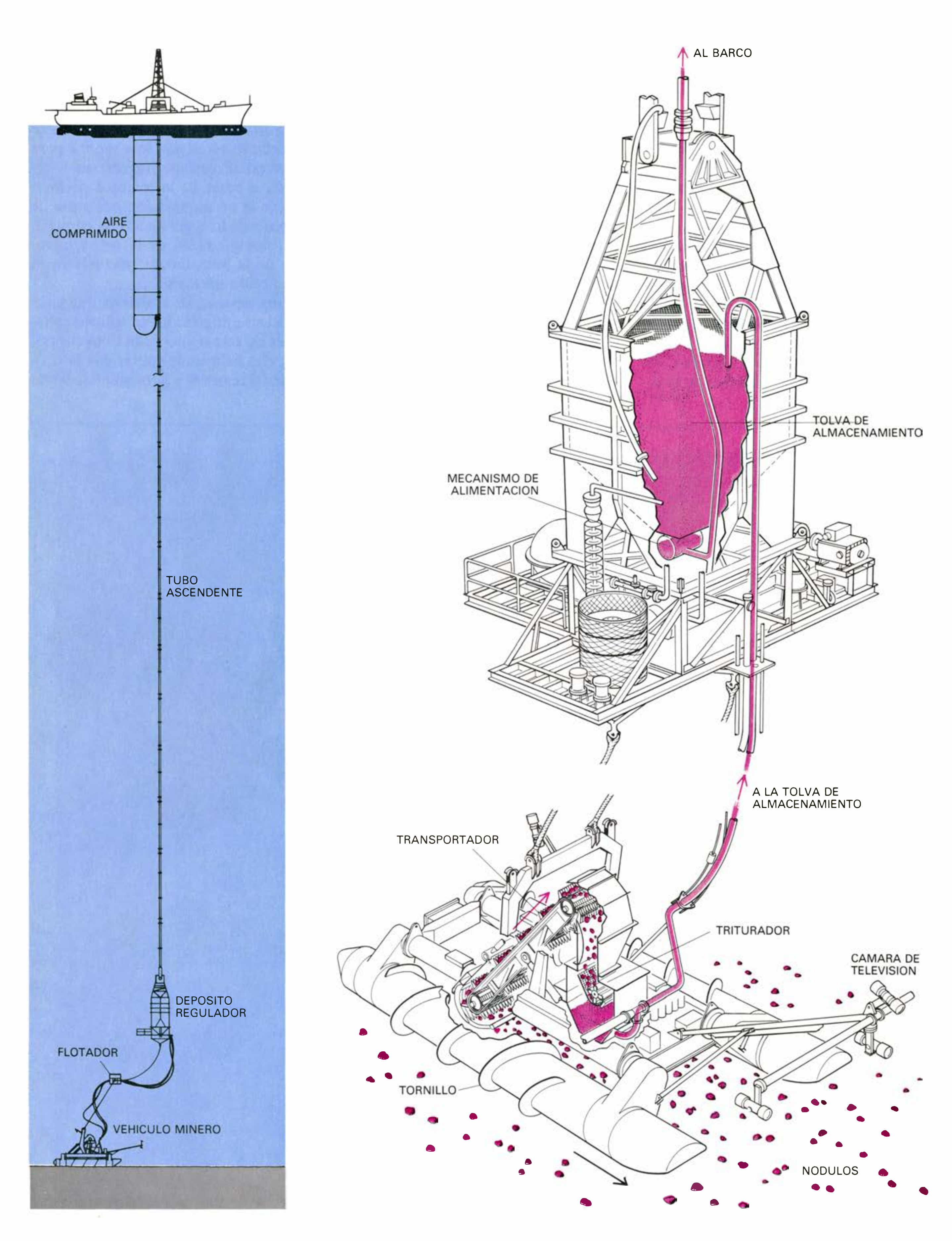
La creación de la zona económica responde también a la necesidad de administrar el medio marino, al tiempo que proporciona un incentivo para el cuidado del mismo. Ello no satisface plenamente las necesidades, pues ni la contaminación ni la pesca respetan los límites nacionales.

En su aspecto negativo, la noción de zona económica se reputa como el mayor robo territorial de la historia. La expansión de la jurisdicción nacional en los océanos aumenta las diferencias entre las naciones. Si alguien esperaba que la Conferencia definiera sin ambigüedades los límites de la jurisdicción nacional, habrá que decir que ha fallado en este aspecto. Los límites de que habla son elásticos, y en cualquier momento pueden surgir reclamaciones. El descubrimiento de cualquier recurso importante en el océano servirá para que el estado costero más cercano lo reclame, a pesar de la solemne declaración de la no apropiación por parte de ningún estado o persona (una corporación, por ejemplo) de la herencia común de la humanidad "más allá de la jurisdicción nacional".

Cabe esperar, sin embargo, que en el futuro se mitiguen las injusticias inherentes en este punto de la Convención. Para ello habrá que contar con la cooperación regional y la consolidación de



marina. Las áreas en naranja son "mares territoriales", para los que se pide derechos de soberanía, sujetos al derecho de "tránsito inocente" (tránsito sin significación militar o económica). Estas peticiones, que afectan a un 40 por ciento de los océanos, se reconocen en la Convención sobre la Ley del Mar. El restante 60 por ciento queda para la tradicional libertad de los mares; los recursos de sus fondos pertenecen a la herencia común de la humanidad.



3. NODULOS DE MANGANESO, expandidos en amplias zonas de los fondos oceánicos, que pueden recuperarse por diversos procedimientos. El sistema que se muestra es el diseñado por Conrad G. Welling y sus colaboradores, de la Ocean Minerals Company. Los nódulos suelen contener un 25 por ciento de manganeso, 1,5 por ciento de níquel, 1,2 por ciento de cobre y 0,2 por ciento de cobalto. En este sistema, el vehículo minero (*abajo, a la derecha*) está propulsado por dos grandes tornillos. Los nódulos se recogen del fondo por una especie de rastrillo sin fin que tritura los nódulos y los envía hasta un depósito regulador (*arriba, a la derecha*) suspendido, por un tubo, del barco. El triturado se bombea luego hasta el barco. Ilustración de Walken Graphics.

las organizaciones internacionales. Como veremos, hay razones para esperar que se alcance un reparto de los beneficios de la explotación de la zona económica.

Los bloques políticos eran algo diferentes en el primer comité, encargado de establecer la gestión de la herencia común de la humanidad. Aquí, las diferencias de intereses entre las naciones industrializadas del norte y las naciones en vías de desarrollo del sur polarizaron el desarrollo de las discusiones. El sur mantenía que la adopción de la Declaración de Principios de la Asamblea General de las Naciones Unidas, de 1970, había sancionado la herencia común de la humanidad con carácter de ley internacional. El primer comité tenía por misión crear una fuerte autoridad sobre los fondos marinos con plenos poderes y funciones. La autoridad dispondría de un brazo ejecutivo, la "empresa", que exploraría y explotaría por sí misma la herencia común de la humanidad en nombre de todos y podría exigir a las compañías mineras la venta de la tecnología necesaria para realizar su labor. Frente a estas propuestas, el norte defendía una autoridad que despachara licencias y cobrara impuestos por explotación de la herencia común, que, por lo demás, quedaba libre para que el norte la utilizara según su criterio.

El compromiso que resolvió este conflicto, costoso, tardó en llegar. El sistema de "empresa" propuesto por el sur no lo aceptaba, por inviable, el norte, seguro de la posesión de la tecnología y el capital necesarios. La idea de las autorizaciones defendida por el norte la rechazaban la mayoría de las naciones presentes en la Conferencia, que veían en ello una violación del principio de herencia común de la humanidad. La solución de compromiso preparada por Henry Kissinger combinaba la inviable "empresa" con las inaceptables autorizaciones en un "sistema paralelo", una combinación que, por la magia de las matemáticas, ¡había de dar una solución viable y aceptable! La Conferencia nunca llegó a recobrarse totalmente de ello.

En el curso de las tediosas negociaciones en las que se puntualizaba el compromiso, el texto del sistema paralelo se iba haciendo cada vez más abstruso y alejado de la realidad. Jugando con la amenaza de la no ratificación por parte del Senado de los Estados Unidos, el negociador norteamericano, Leigh Ratiner, logró que se detallaran por adelantado todos los pormenores administrativos y financieros, fijándolos por escrito en artículos que no podrían alterarse en 25 años –y ello para una industria todavía inexistente y cuyo futuro se presumía bastante incierto. El sur comenzaba ya a replantearse la conveniencia de explotar la herencia común. El descenso de los precios de los productos hizo temer a los países en vías de desarrollo exportadores de recursos la aparición de una nueva competencia para sus materias primas. Contra la oposición del norte (con la excepción de Canadá) estos países se interesaron más por limitar la producción de los fondos marinos que por su administración.

El valor de la herencia común de la humanidad, principalmente de los nódulos de manganeso, ha disminuido ante el desarrollo de otro tipo de explotaciones. México y Chile han encontrado ricos filones de nódulos en sus zonas económicas; Francia tiene nódulos en las aguas de la Polinesia y los Estados Unidos pueden reclamar los que se encuentran cerca de las islas Hawai y otras islas del Pacífico, en cientos de miles de kilómetros cuadrados de fondo oceánico que probablemente tengan nódulos comercialmente explotables. Los nódulos han perdido atractivo ante los depósitos sulfurosos polimetálicos que contienen cobre y otros metales en altas concentraciones, encontrados por primera vez en las fumarolas volcánicas del cinturón del Pacífico, cerca de las islas Galápagos, y frente a las costas de Oregón y Washington. Probablemente se hallen depósitos similares a lo largo de los 65.000 kilómetros de los cinturones volcánicos submarinos de todo el globo.

A pesar de todo, la creación de una autoridad internacional sobre los fondos marinos, aceptada con el consenso de prácticamente la totalidad de la comunidad mundial (debe contarse con el apoyo de los Estados Unidos, ya que se lo dio a la Conferencia y, estoy segura, volverá a darlo), constituye un paso adelante en las relaciones internacionales. Se trata de una institución internacional a la que se ha dotado de poderes para regular y actuar en el marco de la nueva idea de una herencia común de la humanidad. Asimismo, es un primer intento de una política de producción global que no dará la espalda a las exigencias de la conservación del medio. También está abierta a una cooperación industrial entre el norte y el sur, que no se apoye tanto en la ayuda cuanto en el trabajo en común. La Comunidad Económica Europea era, 50 años atrás, el sueño utópico de un puñado de "federalistas". La autoridad internacional sobre los fondos marinos, otro deseo inalcanzable hace 20 años, tiene rango hoy de ley internacional. Algo se ha adelantado.

El tercer comité tenía encomendado determinar un nuevo régimen para la investigación científica marina, para la protección y conservación del medio y la mejora de la transferencia de tecnología. La polarización de las partes antagónicas fue menos crispada en estas deliberaciones; el comité concluyó su tarea antes que los otros.

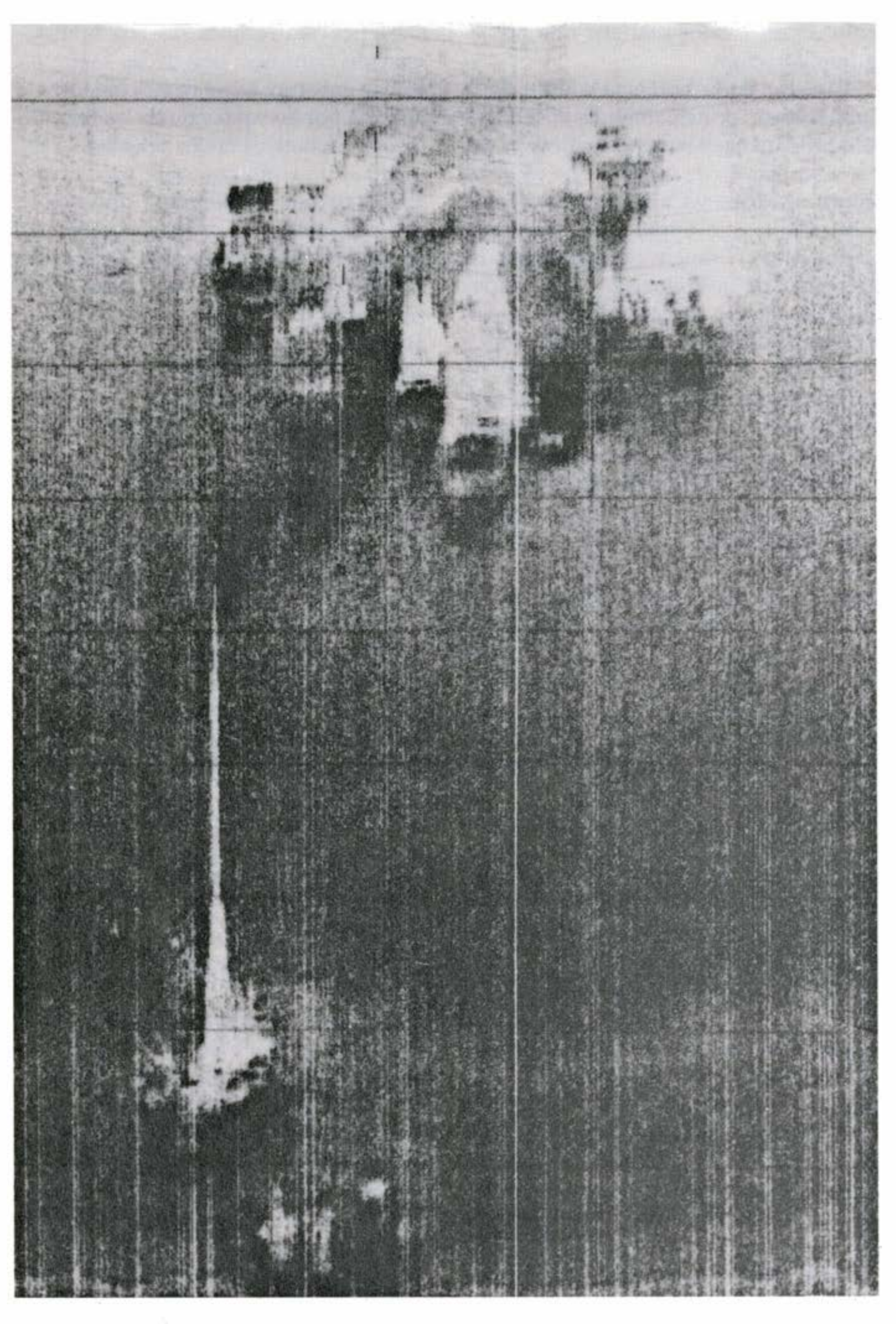
La libertad reconocida para la investigación científica, una de las supuestas en ese entorno, se ve drásticamente recortada por la expansión de las jurisdicciones nacionales en los océanos. Cabe señalar, sin embargo, que, bajo el "régimen de consentimiento" previsto en la Convención, el estado ribereño no negará el permiso a otros países que quieran investigar en sus aguas. (Los científicos norteamericanos que se han quejado de las trabas burocráticas que acompañan a la obtención del consentimiento por parte de un país costero deberían preocuparse por el rechazo de la Convención por los Estados Unidos, pues ello impediría la petición de permisos, no digamos ya su obtención.) Una práctica de consentimiento aumentaría la cooperación bilateral en el campo de la investigación, ajustándola a los intereses de los estados costeros y de los países empeñados en el estudio. Más que cualquier instrumento legal del pasado, la Convención potenciaría la internacionalización de la investigación marina, reiterando la obligación de la cooperación y exigiendo "organizaciones internacionales competentes". Los proyectos amparados en la tutela de esas organizaciones recibirían el consentimiento del país costero, con la única condición de que éste participara en el proyecto y no pusiera trabas al mismo.

Las medidas para la protección y conservación del medio marino previstas en la Convención cubren la contaminación arrastrada por las aguas, la contaminación provocada por la minería submarina y por la producción petrolífera, la contaminación de origen terrestre y la atmosférica. Este marco amplio, sujeto a un código legal internacional, abarca las posibles alteraciones del medio a través de la acuicultura de especies exóticas o a través de la aplicación de nuevas tecnologías. Los estados serán los responsables del daño causado, no sólo en sus aguas territoriales, sino también en mar abierto.

Los organismos competentes, en particular la Organización Marítima Intergubernamental y el Programa de las Naciones Unidas para el Medio Ambiente, unen sus esfuerzos en el trabajo que ahora les ocupa, y que a la vez refuerza su autoridad.

A través de sus proyectos regionales, el Programa de las Naciones Unidas para el Medio Ambiente ha comenzado la transformación de una ley "blanda" de la Convención en una ley "dura", que debe cumplirse. Hay proyectos de esos en marcha, en los que participan 110 estados y gran número de organizaciones intergubernamentales y no gubernamentales [*véase la figura 5*]. Las naciones participantes asumen los gastos de limpieza de los mares y prevención de contaminaciones futuras. Mientras la Conferencia estaba aún reunida, el Programa de las Naciones Unidas para el Medio Ambiente establecía una serie de medidas y normas para el control de la contaminación contempladas en los borradores de la Convención.

Enfrentándose a la indiferencia, más que a la oposición, por parte del norte, los países en vías de desarrollo redactaron los artículos de la Convención sobre la transferencia de tecnología marina. Aunque avanzaron cuanto pudieron, sólo consiguieron leyes "blandas". La transferencia de tecnología supone una compleja labor técnica, social, política y educativa. Cabe preguntarse si realmente puede ajustarse a legislación. La ley establece cuáles pueden ser los objetivos y formas específicas de cooperación: el resto cumple a la historia. Tal historia se está esbozando ya. En los países industrializados, el sistema de patentes imperecederas tiene escasas conexiones con la realidad de una competencia tecnológica. La propiedad intelectual debe encontrar su lugar en la herencia común de la humanidad.

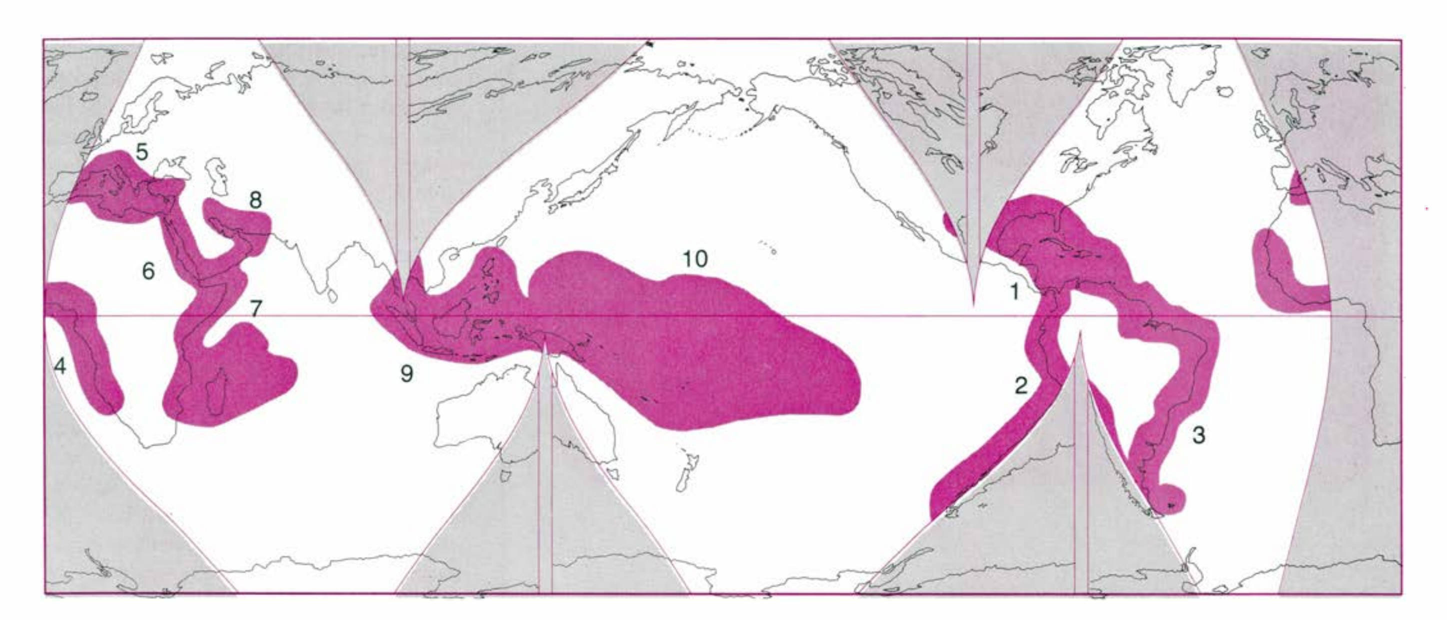


4. DEPOSITOS SUBMARINOS DE MINERALES en el Golfo de California, detectados en una imagen de sonar lateral. Las siluetas que se proyectan del fondo son montículos y "agujas" típicas de áreas de fumarolas volcánicas submarinas próximas a centros de expansión del fondo. En esas áreas, los sulfuros ricos en metales precipitan de la corriente cálida que surge del suelo. Las siluetas sobresalen unos 30 metros del fondo. La distancia entre cuatro líneas horizontales es de 450 metros. La imagen ha sido proporcionada por Fred N. Spiess y Peter F. Lonsdale, de la Institución Oceanográfica Scripps.

La Convención contribuye, de una forma notable, al desarrollo de sistemas para la resolución de disputas y, con ello, al mantenimiento de la paz. De hecho, establece el sistema más amplio y obligatorio de los ideados hasta ahora. Se establece una corte suprema en el Tribunal Internacional de la Ley del Mar, cuya sede se instalará en la ciudad hanseática de Hamburgo. En una de las salas especiales, la de los Fondos Marinos, se dará entrada a contenciosos planteados por corporaciones o por individuos. Inevitablemente, habrá fallos en el sistema. Los estados pueden pedir tratos de favor o excepciones, principalmente en los puntos más sensibles: la delimitación de aguas territoriales o utilización del mar con fines militares.

En esto, así como en otros asuntos, la Convención es un documento ambiguo. La Ley del Mar, sin embargo, no vive una existencia autónoma. Constituye la expresión de la armonía y la tirantez de las relaciones intranacionales e internacionales de nuestro tiempo. Si persisten estas tendencias en otros sectores de las relaciones internacionales, los países explotarán los defectos, fallos y contradicciones de la Convención, a fin de saquear y contaminar los océanos, agotar sus recursos y desplegar sus armas en los fondos marinos. Pero si los estados se inclinan por la negociación de un desarrollo global, equilibrado y por la paz, hallarán en la Convención una base para construir un mundo mejor.

Los océanos están destinados a desempeñar un papel cada vez más importante en la economía mundial. A partir del estado embrionario en que está la acuicultura, la tecnología tradicional de la caza y recolección dará pa-



5. "PLANES DE ACCION" en marcha o en diferentes estados de preparación bajo la dirección de los proyectos marinos regionales del Programa de las Naciones Unidas para el Medio Ambiente. Los planes que se están ejecutando son: (*5*) mar Mediterráneo, (*6*) mar Rojo y golfo de Adén y (*8*) plan de acción regional de Kuwait. Los planes en proyecto son: (*1*) Caribe, (*2*) Pacífico sudoriental, (*3*) Atlántico sudoccidental, (*4*) Africa occidental, (*7*) Africa oriental, (*9*) mares orientales de Asia y (*10*) Pacífico sudoccidental. La proyección es aquí la misma que la del mapa de la figura 2, diseñada por F. W. McBryde.

so al cultivo de plantas acuáticas y a la cría de animales marinos. Los fondos marinos cubrirán las demandas de metales mucho después de que el empobrecimiento de las minas continentales haya agotado las posibilidades de la técnica. A finales de siglo, los fondos marinos aportarán más de la mitad del petróleo mundial y los océanos tendrán reservas de energía en sus mareas, olas y corrientes, en sus gradientes térmicos y salinos, en su biomasa; sin olvidar, ni dejar en último lugar, la energía reservada en los átomos de deuterio contenidos en las moléculas de agua.

La creciente importancia económica del océano se hará sentir directamente en la vida política de las naciones y de la comunidad internacional. Se están creando ministerios para el desarrollo de los océanos, departamentos que asumen nuevas funciones en el control de los mares. Los países están poniendo al día sus leyes marítimas, ampliando la norma hasta aspectos que no contemplaba la legislación. Armonizan su práctica y las instituciones con la Convención de la Ley del Mar y acomodan y completan, de acuerdo con su idiosincrasia, las resoluciones de la Convención.

Los organismos internacionales competentes invocados en la Convención –la Organización Marítima Intergubernamental, la Organización de las Naciones Unidas para la Agricultura y Alimentación, la Comisión Intergubernamental Oceanográfica de la Unesco y el Programa de las Naciones Unidas para el Medio Ambiente– están responsabilizándose de las tareas que les han sido asignadas. En particular, en los estados costeros en vías de desarrollo, estos organismos se encargarán de prestar su ayuda a la elaboración de un marco jurídico nacional, mejorar el control de los recursos naturales en su zona económica, definir las rutas marinas y los esquemas de control del tráfico marino, montar centros regionales de investigación científica y tecnológica y arbitrar en los conflictos.

Aunque parezca paradójico, la expansión de la jurisdicción nacional en los océanos no disminuye la necesidad de la cooperación internacional: por el contrario, la aumenta. No puede haber cooperación internacional si las naciones no están bien organizadas. La realidad tecnológica y económica de nuestros días impone una estrecha cooperación internacional si se quiere que los países no se hundan o quiebren.

Las conexiones a escala mundial de los Programas Regionales del Mar están abriendo las puertas a nuevas formas de cooperación internacional. Tan importante es el interés por el mar, que hasta turcos y griegos, israelitas y árabes, iraquíes e iraníes han continuado su cooperación en los Programas Regionales incluso en tiempo de guerra. Los éxitos del Programa Regional del Océano Indico podrían impulsar la idea de un Mar de la Paz, adoptada y reiterada para aquella región por las Naciones Unidas, que las superpotencias no secundan. En el Mediterráneo, el desarrollo de los mecanismos para llevar a cabo los Programas Regionales quizá resucite la idea presentada por Malta y México en las deliberaciones del Comité de Fondos Marinos en 1972: que las naciones paguen por el uso del océano, por los beneficios obtenidos del petróleo recobrado al mar, por las pesquerías comerciales y por el comercio marítimo.

La firma de la Convención en Montego Bay preveía, para marzo, la puesta en marcha de una nueva organización, la Comisión Preparatoria, encargada de establecer las reglas y regulaciones para la autoridad internacional de los fondos marinos, así como la agenda provisional para la primera reunión de sus órganos de gobierno. La convocatoria de la primera sesión ha de esperar la ratificación de la Convención. En el ínterin, cuya duración no puede preverse, la comisión preparatoria está facultada por la Conferencia para ejercer algunos poderes de la propia autoridad. En este sentido, a la comisión compete examinar las peticiones de los "inversores pioneros" sobre áreas de exploración; verificar los datos de las prospecciones y elegir las sedes para la futura autoridad y su "empresa", y solucionar los problemas de transferencia de tecnología a la autoridad y de preparación de personal en los países en vías de desarrollo. En suma, la comisión preparatoria goza de considerable capacidad operativa. Viene a establecer un régimen interino para la explotación, investigación y desarrollo de la tecnología minera en los océanos, y es probable que, hasta finales de siglo, las actividades de la autoridad de los fondos marinos se limiten a la exploración, investigación y desarrollo.

Una vez adquirida la tecnología necesaria –posiblemente a través de empresas conjuntas o con compañías mineras– la comisión podría aplicarla para explorar los nódulos en las áreas internacionales y ayudar a los países en vías de desarrollo en las prospecciones de los recursos localizados en sus propias zonas económicas. La Convención no rechaza que la autoridad, o la comisión en su caso, haga frente a tales necesidades, siempre que las haya solicitado un país costero.

Otro servicio inmediato y útil que prestaría la comisión sería la creación de parques, regiones libres de explotación y reservas para fines científicos. El primero de esos parques podría ser el cordón volcánico de las Galápagos y de sus fantásticas colonias de gusanos gigantes y crustáceos, sostenidas por las capas de bacterias quimiosintéticas que utilizan azufre en vez de oxígeno, como es común en la vida fotosintética que nos es familiar.

Un tercer servicio inmediato podría relacionarse con el control de armamentos y el desarme. El tratado de 1972 sobre prohibición de utilización de armas nucleares en los océanos no prevé el control de su cumplimiento. Por error, la inspección se dejó al cuidado de las superpotencias. La mayoría de los países pretendían internacionalizar el control y seguimiento, pero hasta ahora sus deseos se han visto frustrados por la falta de un organismo internacional competente. La comisión preparatoria y la autoridad podrían cubrir esa laguna.

En 1987, cuando se revise el tratado de desarme de los fondos marinos, la Comisión puede muy bien haber adquirido los conocimientos necesarios; la tecnología acústica, sísmica, magnética y electrónica empleada en la exploración de recursos podrá aplicarse entonces a los trabajos de inspección. Una enmienda al tratado, en 1987, que apoyaran la mayoría de los países firmantes, podría confiar esta tarea a la comisión preparatoria, o a la propia autoridad si la Convención se hubiera ratificado.

La firma de la Convención, en Montego Bay, por la aplastante mayoría de la comunidad internacional abre una nueva fase en la construcción de un orden nuevo. La Convención ofrece nuevas formas de intensificar el desarrollo económico, el desarme y la protección del medio ambiente y la creación de nuevas formas de cooperación científica e industrial entre el norte y el sur. Todas esas posibilidades no se darían sin la nueva Ley del Mar.

Microprogramación

En casi todos los ordenadores modernos, el flujo de información está controlado en último extremo por un microprograma. Por tal se entiende un sistema de instrucciones que sustituye a los sistemas de control por medios circuitales

David A. Patterson

Cuestión fundamental al diseñar cualquier ordenador electrónico es el control, o pilotaje, de las señales eléctricas portadoras de información. En la unidad lógico-aritmética, donde realmente tiene lugar el procesado de las informaciones, es preciso ir encauzando y dirigiendo las señales entre diversos contadores, sumadores y demás componentes. El sistema de control debe intermediar igualmente en la transferencia de información entre el procesador central, las unidades de memoria principal y los distintos dispositivos de entrada y salida de datos. Una de las soluciones consiste en implantar permanentemente el sistema de control en forma de circuitos eléctricos que sean parte integrante del procesador. Hay una segunda solución, más versátil, y en muchos casos menos onerosa. Se trata, en definitiva, de reducir la complejidad del sistema de control registrando la lista detallada de instrucciones destinadas al control del ordenador en forma codificada. Dicho con otras palabras, la secuencia de rutas que ha de ir siguiendo una señal queda incorporada en un programa, que se almacena en una unidad individual de memoria con la que será provisto el procesador.

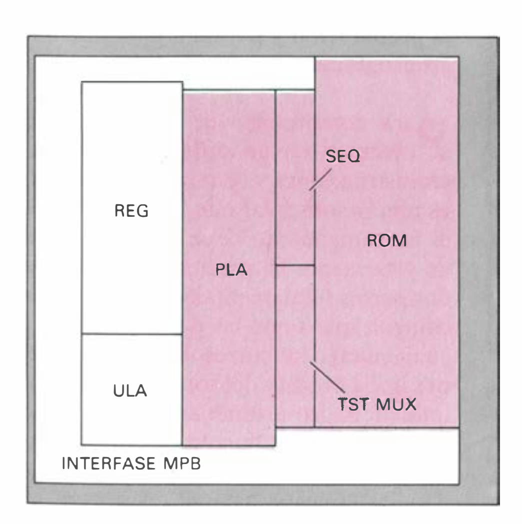
Dentro del sistema jerárquico de los programas que gobiernan y hacen funcionar un ordenador, las instrucciones ejecutadas por el sistema de control ocupan el nivel más elemental y último, por abajo; cada una de sus instrucciones especifica un estado funcional de la máquina. Siendo las instrucciones de control responsables de tan minuciosos detalles, la tarea de definirlas y codificarlas ha sido bautizada "microprogramación", diferenciándola así de la redacción y puesta a punto de programas de más alto nivel, que en su conjunto forman el soporte lógico (*software*) del ordenador. Se dice, además, que los sistemas de instrucciones de control –los microprogramas– están escritos en microcódigo. La microprogramación se ideó hace más de 30 años, al poco del advenimiento de los primeros ordenadores. En aquellos tiempos no se disponía de soporte físico (*hardware*) para materializar y llevar a efecto tal idea. Este es, sin embargo, el método que se ha adoptado en la mayoría de los ordenadores que se construyen en nuestros días.

No poca confusión viene reinando estos últimos años acerca del significado del término microprogramación, a causa, sobre todo, del advenimiento de los microprocesadores, las "computadoras de pastilla" (*chip*) que se encuentran en el corazón de los últimos productos de la miniaturización progresiva, consecuencia de los avances tecnológicos en semiconductores de silicio. Es preciso insistir en que no es lo mismo la microprogramación de ordenadores que la programación de microordenadores; en principio, todo ordenador, desde el mayor de los "sistemas maestros" (*mainframe systems*) hasta el más pequeño de los ordenadores personales, puede ser diseñado con sistemas de control microprogramado. Para evitar semejantes confusiones, está imponiéndose la expresión soporte fijo (*firmware*) para designar a los microprogramas, pretendiendo así expresar su posición intermedia entre la organización electromecánica del ordenador (*hardware*) y sus sistemas de programas (*software*). Para mayor claridad, adoptaremos aquí el siguiente convenio: el término macroinstrucción aludirá a instrucciones de soporte lógico, mientras que microinstrucción hará referencia a instrucciones de soporte fijo (*firmware*).

Antes de abordar la naturaleza de la microprogramación será necesario recordar algunos principios básicos de la configuración material y de la programación de ordenadores. Un ordenador es, en muchos aspectos, semejante a una simple calculadora de bolsillo. En ambos casos, la introducción de números se realiza por medio de un teclado y la lectura de resultados se efectúa sobre alguna unidad de presentación, por lo común una pantalla. Muchas calculadoras disponen de una tecla de memoria para registrar números que más tarde podrán recuperarse; también la memoria de un ordenador tiene la función de almacenar y luego permitir la recuperación de números. Empero, la memoria de un ordenador es capaz de manejar millones de números. Para poder llevar la gestión de tantísimos números a cada uno se le asigna una "dirección" específica.

Tanto a calculadoras como a ordenadores hay que decirles lo que tienen que hacer. Dar una instrucción a un ordenador viene a ser como pulsar una de las teclas de carácter aritmético de la calculadora. Por ejemplo, casi todos los ordenadores admiten las macroinstrucciones ADD, SUB, MULT y DIV, que equivalen exactamente a las teclas aritméticas +, −, ×, y ÷ de las calculadoras. La macroinstrucción LOAD desempeña un papel análogo a la tecla de memoria de una calculadora.

Mas, a diferencia de las calculadoras sencillas, los ordenadores son, ante todo, máquinas que toman decisiones. Para comprender en qué consiste la diferencia, fijémonos en la sucesión de operaciones representadas en las últimas secciones del Modelo 1040, impreso de la declaración de impuesto sobre la renta de los ciudadanos norteamericanos [*véase la figura 3*]. Un programador de ordenadores que tuviera la tarea de cumplimentar este formulario comenzaría preparando un diagrama de flujo (llamado también flujograma) que le sirve a modo de anteproyecto de construcción del programa [*véase la figura 4*]. Un contribuyente, trabajando con una sencilla calculadora, podría ir siguiendo el flujograma, ejecutando ordenadamente las instrucciones, hasta llegar al paso "de decisión", enmarcado en un rombo. El contribuyente tendría entonces que comparar los números



1. UN MICROPROCESADOR MICROPROGRAMADO ocupa el corazón de los ordenadores de la familia HP-9000, de la Hewlett-Packard Company. El circuito microelectrónico aquí fotografiado es de la máxima densidad alcanzada en este tipo; sobre una pastilla (*chip*) cuadrada de 6,3 milímetros de lado se han inscrito el equivalente de 450.000 transistores. La separación típica entre líneas de transporte de señales es de un micrometro. El diagrama de la izquierda permite distinguir las siete subsecciones principales. La zona coloreada distingue a las subsecciones constituyentes del sistema de control del procesador; la zona en blanco corresponde a la senda de datos. El sistema de control consta de cuatro partes: una memoria de sólo lectura (*ROM*), donde se encuentran almacenadas las microinstrucciones encargadas de gestionar el flujo de información en la totalidad del ordenador; una configuración de lógica programada (*programmed-logic array*, *PLA*), donde se procede a decodificar la mayor parte de las microinstrucciones antes de que sean transmitidas a sus diferentes destinos; un dispositivo secuenciador (*SEQ*) en donde se prepara el orden de las diversas microinstrucciones; por último, un multiplexador-comprobador (*TST MUX*), donde se decodifican ciertas microinstrucciones condicionales. Las otras tres subsecciones son: una formación de registros de memoria (*REG*) donde los datos que está manipulando el ordenador en ese instante son almacenados en forma de "palabras", cada una de 32 bits de longitud; una unidad lógico-aritmética (*ULA*), donde verdaderamente se lleva a cabo el procesado de estas palabras de 32 bits; y una interfase memoria-procesador-bus, encargada de supervisar las comunicaciones entre el microprocesador y el resto del ordenador. Las formaciones rectangulares situadas en los márgenes superior, derecho e inferior del microcircuito corresponden a dispositivos de verificación, patrones de resolución especial y marcas de alineación.

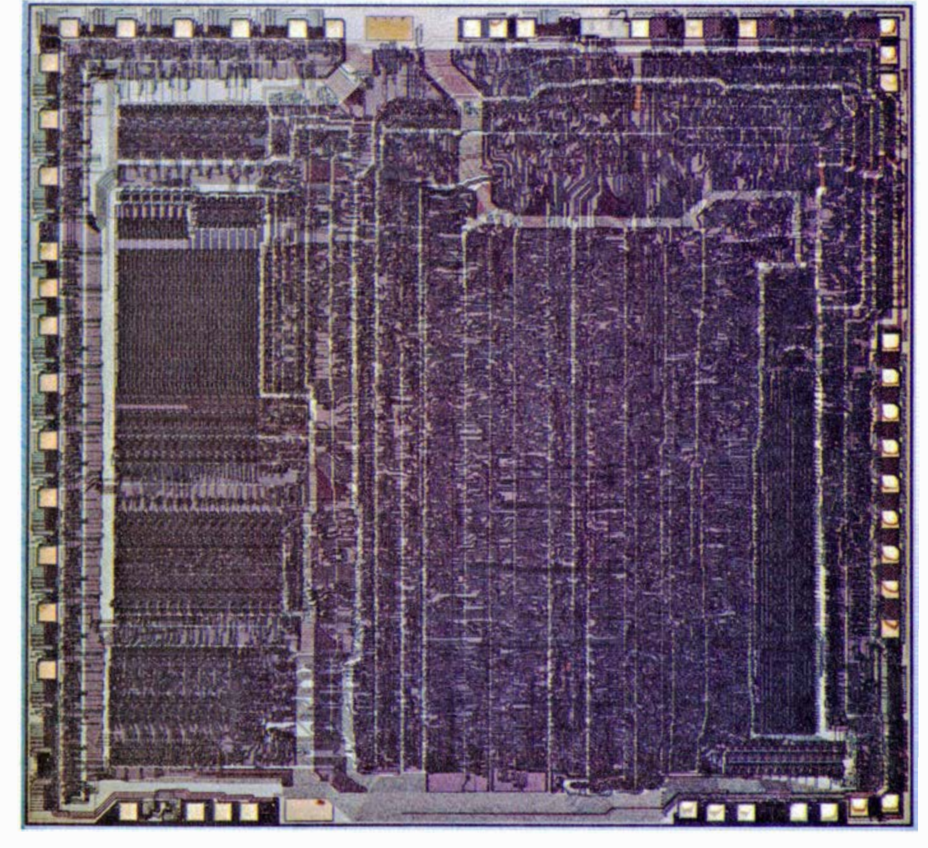
contenidos en las dos cajas precedentes y, según cuál de ellos sea mayor, tomar la ruta correspondiente y proseguir el cálculo.

Si el contribuyente, en lugar de una calculadora, dispusiera de un ordenador, nuestro programador podría ir un paso más allá: hacer que el ordenador tomase la decisión por sí mismo. De acuerdo con el diagrama de flujo, empezaría instruyendo al ordenador para que sumase las cantidades inscritas en las líneas números 50 a 58 del impreso de declaración, almacenando el total de cada adición en una dirección específica de la memoria del ordenador; en este caso, las direcciones de los números han sido designadas *L*50, *L*51, *L*52, y así sucesivamente, y van sumándose, por turno, a una dirección llamada "acumulador" y designada por *A* [*véase la figura 5*]. El total correspondiente a esta sucesión de pasos se almacenará entonces en la dirección *L*59. A continuación, el programador haría otro tanto con las líneas 60 a 66, almacenando el total de la segunda sucesión en la dirección *L*67. En este punto (paso 19), el programador daría la siguiente macroinstrucción: "JUMPIF *A* < *L*59, 23". Esta macroinstrucción, llamada "salto condicional", dice a la computadora que comience la ejecución de una nueva secuencia de macroinstrucciones que arranca del paso 23 si es que el número alojado en la dirección usada como acumulador es menor que el alojado en la dirección *L*59. En otro caso, el programa continuaría procediendo a ejecutar la instrucción número 20.

La macroinstrucción de salto condicional, correspondiente al recuadro rómbico del flujograma, confiere al ordenador capacidad para tomar decisiones, suponiendo, claro está, que el programador consiga representar cada decisión por medio de comparaciones aritméticas sencillas. El programador construye un programa tejiendo una red de saltos condicionales que vayan conduciendo a cada uno de los datos de entrada hasta su correcto destino. Por término medio, en los programas de ordenador una de cada diez macroinstrucciones es una macroinstrucción de salto condicional; muchos de los ordenadores hoy disponibles pueden ejecutar una macroinstrucción en menos de una millonésima de segundo; un ordenador efectuará, pues, unas 100.000 decisiones de tipo semejante al anterior cada segundo. Las macroinstrucciones correspondientes a los procesos de decisión están a su vez codificadas en clave numérica y almacenadas, lo mismo que los demás números que el ordenador ha de manejar, en la memoria principal del mismo. El principio de programa almacenado, cuya invención ha sido piedra angular en el desarrollo de los modernos ordenadores, permite cambiar la función desempeñada por la máquina con sólo modificar el contenido de su unidad de memoria, sin alteración alguna en sus circuitos eléctricos.

Conforme los ordenadores van adquiriendo una rapidez mayor, más macroinstrucciones pueden ejecutar, y mayor número de decisiones tomar, en un mismo intervalo de tiempo; resulta posible, por tanto, hacer funcionar en ellos programas de complejidad cada vez mayor. En tales circunstancias resulta difícil evitar la comisión de errores. Los programadores prefieren un término cuyas connotaciones no son exactamente las mismas que las de error: dicen más bien que el programa tiene "gazapos" (*bugs*). Para mejor comprender lo que con este término pretenden expresar, volvamos a fijarnos en la ilustración donde se reproduce el impreso de declaración de renta de las personas físicas [*figura 3*]. La línea 68 le dice al declarante qué debe hacer si la cantidad inscrita en la línea 67 es mayor que la anotada en la línea 59, y la línea 71 explica cómo proseguir si la cifra de la línea 59 es mayor que la de la línea 67; ¡pero en ningún lugar del impreso se le aclara cómo proceder si las cantidades de las líneas 59 y 67 resultan iguales! Evidentemente, a la Dirección General de Renta de las Personas Físicas se le ha escapado aquí un "gazapo". Son precisamente los detalles de este tipo a los que el programador debe prestar cuidadosa atención al redactar y "depurar" (*debug*) programas. La dificultad de la tarea se ve multiplicada cuando hay que entrar en los detalles de estructura fina que conlleva la preparación y puesta a punto de microprogramas.

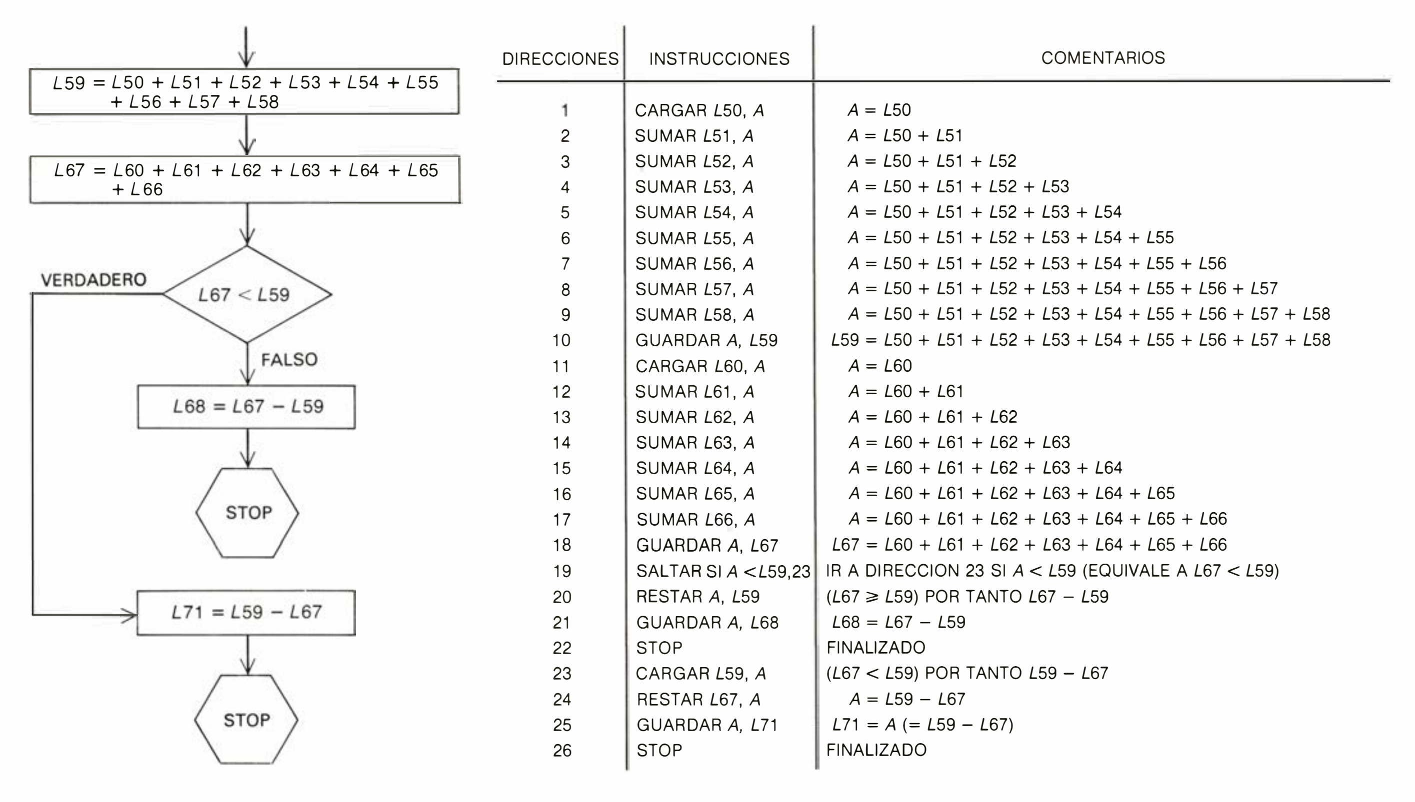
Para comprender de qué modo se ejecutan en un ordenador las macroinstrucciones y se toman decisiones, es preciso observar más detenidamente el funcionamiento de la unidad central de procesado. El procesador consta de dos partes fundamentales: el sistema de control, que toma las decisiones, y del cual emanan las correspondientes órdenes hacia el resto del soporte físico, y la unidad lógico-aritmética, o senda de datos, donde realmente se llevan a cabo los cálculos. En el corazón de la unidad lógico-aritmética se encuentra un



2. SE DISTINGUEN LOS MICROPROCESADORES NO MICROPROGRAMADOS por una sección de control de apariencia irregular, construida e integrada en el circuito, y que ocupa la gran zona de la derecha de la fotografía. Las dos zonas menores y de aspecto más regular situadas a la izquierda corresponden a las subsecciones *REG* y *ULA* de la figura 1. La irregularidad de tales sistemas de control integrados en el circuito dificulta su diseño o posterior modificación. En la época en que se construyó esta pastilla (hará unos cinco años) se creía que la microprogramación exigiría dispositivos mayores.

	Line		Box
	50	**Balance.** Subtract line 49 from line 40 and enter difference (but not less than zero) . ▶	50
Other Taxes	51	Self-employment tax *(attach Schedule SE)*	51
	52	Minimum tax *(attach Form 4625)*	52
(Including Advance EIC Payments)	53	Alternative minimum tax *(attach Form 6251)*	53
	54	Tax from recapture of investment credit *(attach Form 4255)*	54
	55	Social security (FICA) tax on tip income not reported to employer *(attach Form 4137)*	55
	56	Uncollected employee FICA and RRTA tax on tips *(from Form W–2)*	56
	57	Tax on an IRA *(attach Form 5329)*	57
	58	Advance earned income credit (EIC) payments received *(from Form W–2)*	58
06	59	**Total tax.** Add lines 50 through 58 ■	59
Payments	60	Total Federal income tax withheld	60
	61	1982 estimated tax payments and amount applied from 1981 return	61
Attach Forms W–2, W–2G, and W–2P to front.	62	Earned income credit. If line 33 is under $10,000, see page 15 of Instructions	62
	63	Amount paid with Form 4868	63
	64	Excess FICA and RRTA tax withheld (two or more employers)	64
	65	Credit for Federal tax on special fuels and oils *(attach Form 4136)*	65
	66	Regulated Investment Company credit *(attach Form 2439)*	66
	67	**Total.** Add lines 60 through 66 ▶	67
Refund or Amount You Owe	68	If line 67 is larger than line 59, enter amount **OVERPAID** ▶	68
	69	Amount of line 68 to be **REFUNDED TO YOU** ▶	69
	70	Amount of line 68 to be applied to your 1983 estimated tax ▶	70
	71	If line 59 is larger than line 67, enter **AMOUNT YOU OWE.** Attach check or money order for full amount payable to Internal Revenue Service. Write your social security number and "1982 Form 1040" on it. ▶ (Check ▶ ☐ if Form 2210 (2210F) is attached. See page 16 of Instructions.) ▶ $	71

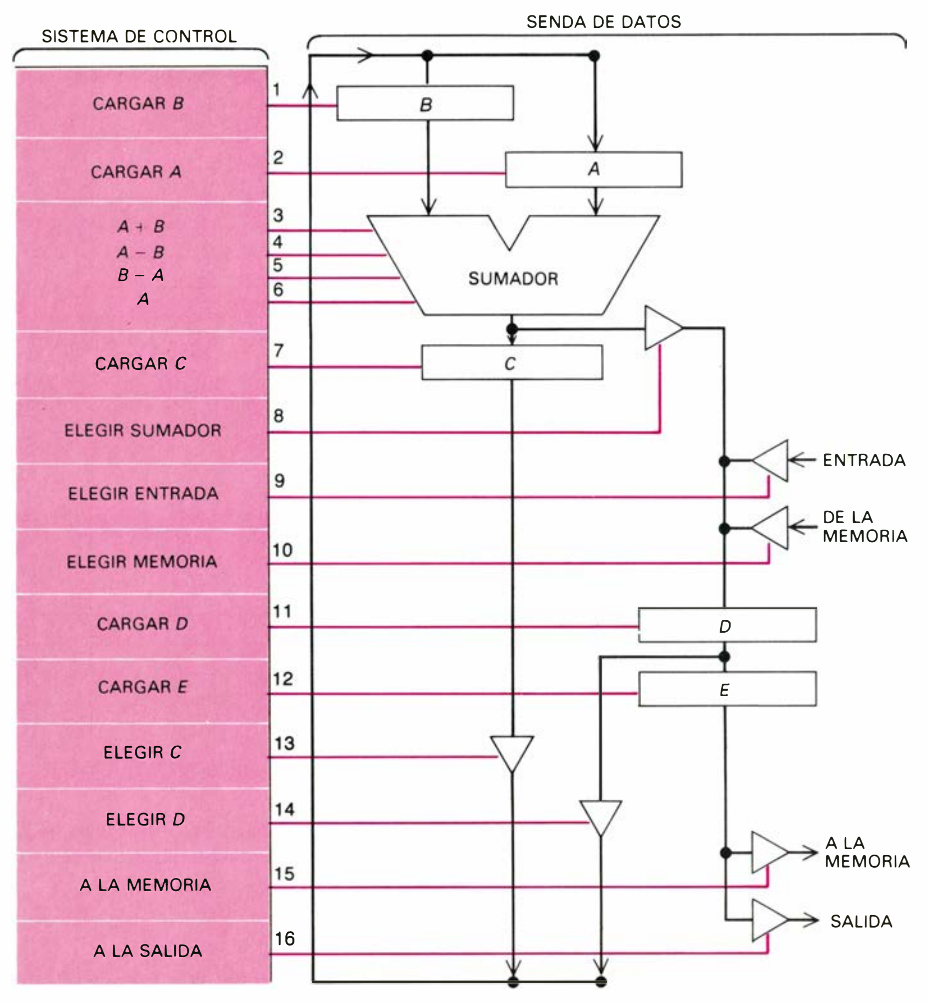
3. UN IMPRESO de declaración de renta norteamericano puede servirnos de ejemplo de secuencia de operaciones ejecutables mediante un programa de ordenador. La ilustración reproduce las 22 últimas líneas del Modelo 1040, para declaración de renta de las personas físicas durante el ejercicio de 1982.



DIRECCIONES	INSTRUCCIONES	COMENTARIOS
1	CARGAR *L*50, *A*	*A* = *L*50
2	SUMAR *L*51, *A*	*A* = *L*50 + *L*51
3	SUMAR *L*52, *A*	*A* = *L*50 + *L*51 + *L*52
4	SUMAR *L*53, *A*	*A* = *L*50 + *L*51 + *L*52 + *L*53
5	SUMAR *L*54, *A*	*A* = *L*50 + *L*51 + *L*52 + *L*53 + *L*54
6	SUMAR *L*55, *A*	*A* = *L*50 + *L*51 + *L*52 + *L*53 + *L*54 + *L*55
7	SUMAR *L*56, *A*	*A* = *L*50 + *L*51 + *L*52 + *L*53 + *L*54 + *L*55 + *L*56
8	SUMAR *L*57, *A*	*A* = *L*50 + *L*51 + *L*52 + *L*53 + *L*54 + *L*55 + *L*56 + *L*57
9	SUMAR *L*58, *A*	*A* = *L*50 + *L*51 + *L*52 + *L*53 + *L*54 + *L*55 + *L*56 + *L*57 + *L*58
10	GUARDAR *A*, *L*59	*L*59 = *L*50 + *L*51 + *L*52 + *L*53 + *L*54 + *L*55 + *L*56 + *L*57 + *L*58
11	CARGAR *L*60, *A*	*A* = *L*60
12	SUMAR *L*61, *A*	*A* = *L*60 + *L*61
13	SUMAR *L*62, *A*	*A* = *L*60 + *L*61 + *L*62
14	SUMAR *L*63, *A*	*A* = *L*60 + *L*61 + *L*62 + *L*63
15	SUMAR *L*64, *A*	*A* = *L*60 + *L*61 + *L*62 + *L*63 + *L*64
16	SUMAR *L*65, *A*	*A* = *L*60 + *L*61 + *L*62 + *L*63 + *L*64 + *L*65
17	SUMAR *L*66, *A*	*A* = *L*60 + *L*61 + *L*62 + *L*63 + *L*64 + *L*65 + *L*66
18	GUARDAR *A*, *L*67	*L*67 = *L*60 + *L*61 + *L*62 + *L*63 + *L*64 + *L*65 + *L*66
19	SALTAR SI *A* <*L*59,23	IR A DIRECCION 23 SI *A* < *L*59 (EQUIVALE A *L*67 < *L*59)
20	RESTAR *A*, *L*59	(*L*67 ≥ *L*59) POR TANTO *L*67 − *L*59
21	GUARDAR *A*, *L*68	*L*68 = *L*67 − *L*59
22	STOP	FINALIZADO
23	CARGAR *L*59, *A*	(*L*67 < *L*59) POR TANTO *L*59 − *L*67
24	RESTAR *L*67, *A*	*A* = *L*59 − *L*67
25	GUARDAR *A*, *L*71	*L*71 = *A* (= *L*59 − *L*67)
26	STOP	FINALIZADO

4. ESTE FLUJOGRAMA sirve para mostrar el plan de organización de la serie de operaciones precisas para cubrir las 22 últimas líneas del impreso de declaración de renta. Las cantidades anotadas en cada línea del impreso son almacenadas en una dirección específica de la memoria del ordenador; en este caso, las direcciones correspondientes a los números se han designado *L*50, *L*51, *L*52, etcétera. La caja rómbica es el símbolo convencional que representa los pasos "de decisión".

5. PROGRAMA DE MACROINSTRUCCIONES para cumplimentar el impreso de la declaración de la renta, deducido del flujograma anterior. Cada macroinstrucción está a su vez almacenada en una dirección de la memoria (*primera columna de la izquierda*). Los números situados en las direcciones *L*50, *L*51, *L*52, etcétera, se van sumando sucesivamente a una dirección "acumulador", la designada *A*. Los pasos 1 a 10 corresponden a la caja situada en lo más alto del flujograma, mientras que los pasos 11 a 18 corresponden a la segunda de las cajas contando desde arriba. La macroinstrucción de salto condicional situada en el paso 19 ("JUMPIF *A* < *L*59, 23") equivale al paso de decisión representado por un rombo en el flujograma. Los pasos 20 y 21 son los correspondientes a la rama "Falso" del flujograma; los pasos 23, 24 y 25 se corresponden con la rama "Verdadero". Los comentarios de la derecha sirven para recordar al programador qué es lo que va aconteciendo; no afectan en absoluto a la ejecución del programa.



6. PROCESADOR CENTRAL IDEALIZADO, que se basa en 16 líneas de control (*trazadas en color*) para ir encauzando a través de la senda de datos simplificada (*en líneas negras*) las señales eléctricas que representan información. En este caso, la unidad lógico-aritmética dispone únicamente de un sistema especializado: un sumador, que recibe números a través de dos líneas de entrada, los suma o resta de acuerdo con las órdenes que reciba a través de las líneas de control *3* a *6* y envía el resultado a través de una línea de salida. Las cajas rectangulares alargadas son registros, dispositivos de memoria para uso inmediato, capaces de alojar una palabra de información cada vez. Los triángulos simbolizan puertas, dispositivos que permiten el paso de información cuando la línea de control a ellos asociada se encuentra excitada, y que la bloquean cuando la línea de control se encuentra en estado inerte. En los microprocesadores no microprogramados cada una de las órdenes enumeradas en la sección de control está definida mediante circuitos permanentes especializados. En la versión microprogramada del mismo procesador, cada orden está almacenada en forma de microinstrucción individual en la memoria de microprograma. En la figura número 8 se ilustra un microprograma de muestra para este procesador simplificado.

sumador, pues en principio todas las funciones aritméticas pueden reducirse a pasos de adición reiterada. Cada sección del procesador dispone de sus propios dispositivos de memoria "a corto plazo", llamados registros.

A este nivel, la organización de un ordenador recuerda una instalación de fontanería. Los cables y pistas conductoras que conectan los diversos circuitos pueden imaginarse como cañerías que encauzan el flujo de información. El sistema de control dirige este flujo enviando señales de dos tipos, uno encargado de abrir y otro de cerrar gran número de dispositivos electrónicos semejantes a válvulas de paso, llamados "puertas" [*véase la figura 8*]. Dado que la apertura y cierre de las válvulas debe producirse en los momentos oportunos para poder encauzar la información hacia los lugares precisos, es necesario que las señales de control estén sincronizadas con gran exactitud. Por este motivo, la cadencia de funcionamiento de todo el procesador está subordinada a un reloj interno. Cada macroinstrucción corresponde a una secuencia de ciclos de reloj, y cada ciclo determina la transferencia de una información individual a lo largo de la senda de datos. La función de sistema de control es suministrar las señales de control durante los ciclos de reloj correspondientes.

En los sistemas de control materializados por medios circuitales es preciso idear una red de lógica electrónica capaz de reconocer e identificar cada una de las macroinstrucciones del repertorio propio del ordenador. En sí misma, la macroinstrucción es una pauta de dígitos binarios materializada en una secuencia de señales dirigidas al procesador. El sistema de control tiene que transformar, mediante algún procedimiento, esta pauta en otro sistema totalmente diferente de señales encargadas de abrir y cerrar diversas puertas situadas en la senda de datos. Además, las señales emitidas por el sistema de control han de estar temporizadas, de modo que coincidan con los oportunos ciclos de reloj. Cada macroinstrucción ha de dar lugar a una secuencia distinta de señales de control, y por ello la complejidad del sistema se hace mayor conforme aumenta el número de posibles macroinstrucciones. La única forma, por otra parte, de introducir cambios o corregir deslices es reformar los circuitos.

En los primeros tiempos, el diseño de los sistemas de control de ordenadores era problema de enorme dificultad. Comparada con él, la tarea de diseñar los restantes elementos mecánicos y circuitales era francamente fácil. Un teletipo podía servir como dispositivo de entrada y salida; y el carácter repetitivo de las estructuras de memoria y de los elementos circuitales de la senda de datos simplificaba su diseño y construcción. En cambio, los circuitos necesarios para llevar la cuenta de los ciclos de reloj y para emitir las señales de control en los instantes oportunos eran causa de que el sistema de control tuviese una composición muy irregular, lo que dificultaba mucho la tarea de diseñar esta parte del procesador central.

En 1949, Maurice V. Wilkes, de la Universidad de Cambridge, se propuso descubrir un procedimiento para organizar mejor la función de control. Tras dos años de estudio, Wilkes llegó a la conclusión de que lo más conveniente era imaginar el sistema de control como una matriz o casillero rectangular, donde cada fila de cuadros correspondería a un ciclo de reloj, y cada columna estuviera asociada con una línea de control. Por ejemplo, el procesador simplificado que vemos en la figura 6 tiene 16 líneas de control y, por tanto, la correspondiente matriz ha de tener 16 cuadros de anchura. Elegir una secuencia de operaciones se convierte entonces en la sencilla cuestión de colocar en los cuadrados la combinación idónea de símbolos binarios: por cada línea de control que deba encontrarse en estado "excitada" durante un determinado ciclo de reloj anotaremos un 1 en la casilla correspondiente, y por cada línea que deba permanecer "desconectada" escribiremos un 0.

La traducción circuital de la matriz de control es una estructura de memo-

ria sencilla y repetitiva. El contenido de cada casilla de cada una de las filas determina el estado en que se encontrará la correspondiente línea de control durante el intervalo de duración de un ciclo de reloj. La pauta de dígitos que componen cada macroinstrucción sirve únicamente para seleccionar una fila (o sucesión de filas) en la memoria de control. Dicho de otra forma, la macroinstrucción se convierte en una dirección que designa la fila. El diseño de un sistema de control pasa así a ser más cuestión de organización lógica (de *software*) que de organización circuital (de *hardware*). La complejidad del sistema de control reside entonces más en la dificultad de elegir las secuencias idóneas de 1*s* y 0*s* que en especificar las combinaciones correctas de circuitos capaces de generar las señales de control.

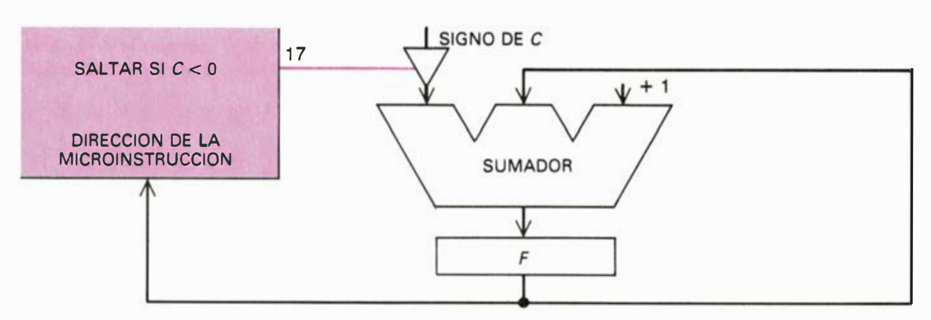
Wilkes supo reconocer las semejanzas existentes entre su solución y la programación ordinaria; para describir sus ideas tomó prestados cierto número de términos del vocabulario de programación ordinaria, anteponiendo en cada caso el prefijo "micro-" para denotar la naturaleza elemental de la función de control. De aquí la voz "microprogramación" y su familia de vocablos asociados. En particular, Wilkes denominó microinstrucción a cada una de las hileras de cuadros de la matriz de control, y llamó microprograma a cada secuencia de hileras encargadas de ejecutar una macroinstrucción. La estructura de memoria diseñada para albergar el microprograma se llama memoria de microprograma. Gracias a la noción de microprograma, la función de control es mucho más sencilla de comprender; al propio tiempo, al sustituir los complejos circuitos de control por una organización repetitiva de celdillas de memoria, simplificó mucho el diseño y la construcción de los circuitos. Y lo que es de la máxima importancia: dio al ordenador incrementada flexibilidad, pues ahora podía modificarse el sistema de control sin necesidad de volver a diseñar el soporte físico.

Fijémonos en el microprograma de muestra correspondiente al procesador esquemático antes considerado [*véase la figura 8*]. Supongamos que la operación controlada por el macroprograma sea restar el número almacenado en el registro de memoria *B* del alojado en el registro acumulador *A*, y almacenar después el resultado en *A*. Inspeccionando la senda de datos en el diagrama circuital podemos ver que el dispositivo etiquetado "sumador" es capaz, no sólo de sumar los dos números, sino también de restar del otro uno cualquiera de ellos. Empero, antes de que el resultado pueda ser devuelto a *A* es preciso albergarlo temporalmente en *C*; por consiguiente, el registro *C* ha de ser cargado primero, y luego descargado, en la senda de datos que conduce de retorno a *A*. La operación de sustracción exige dos microinstrucciones. La primera le dice al sumador que reste el número alojado en *B* del que se encuentra en *A*, y que ponga el resultado en *C*; de acuerdo con esto, se activan las líneas de control números 4 ("$A-B$") y 7 ("CARGAR *C*"). La siguiente microinstrucción transfiere la información desde *C* hasta *A*, abriendo una puerta en la línea que parte de *C* para poder trasladar la información hasta *A*; así pues, las líneas de control números 13 ("SELECCIONAR *C*") y 2 ("CARGAR *A*") quedan activadas.

Examinemos ahora un segundo microprograma, algo más complicado, para este mismo procesador. El segundo microprograma requiere efectuar un proceso de decisión análogo al salto condicional del programa para declaración de renta. La decisión se reduce a elegir una de las dos microinstrucciones. A tal propósito es preciso añadir un cuadrado más a las microinstrucciones y una línea adicional de control al procesador. El soporte físico del dispositivo de control así ampliado puede verse en la figura 7 y, el microprograma revisado, en la 9.

La operación controlada por el segundo microprograma consiste en cargar el número situado en el registro *D* en el registro *E*, si *A* es mayor o igual que *B*. Se resta primero *B* de *A*, y el resultado se almacena en *C*. Si el número alojado en *C* es negativo, se sigue que *A* es menor que *B* y se decidirá entonces pasar por alto la microinstrucción siguiente, reanudando la ejecución del microprograma en la microinstrucción número 4. Por otra parte, si el resultado de la sustracción es un número positivo o cero, *A* es mayor o igual que *B* y el microprograma prosigue con la ejecución de la orden sucesiva inmediata, que carga en *E* el contenido de *D*. Se ha añadido un cuadrado a las microinstrucciones para dar la orden de saltar a la microinstrucción número 4 si *C* es negativo, y al sistema circuital del procesador se añade la correspondiente línea de control.

Se dice que el microcódigo va "interpretando" las macroinstrucciones y va produciendo las señales de control que realmente determinan el flujo de información a lo largo de la senda de datos. Así pues, la secuencia de microinstrucciones del segundo de los ejemplos anteriores es la interpretación en microcódigo de la instrucción condicional que fue denominada "JUMPIF" en el paso número 19 del programa de la declaración de renta. El microprograma se ocupa del sencillo criterio aritmético presentado por el ordenador, ordenando al soporte físico que inspeccione si el número contenido en el registro *C* es negativo y, si lo es, que ejecute la microinstrucción condicional apropiada. Para que los dispositivos circuitales lleven a cabo semejante comprobación tendrán que realimentar en el sistema de control informaciones tomadas de la senda de datos. El circuito eléctrico necesario para determinar si el número alojado en *C* es negativo o no consiste, sencillamente, en un conductor que parte de *C*, y que representa el signo del número contenido en el registro. La línea está desexcitada (estado 0) cuando tal signo es positivo (lo que equivale



7. SE REQUIERE SOPORTE FISICO ADICIONAL para capacitar a la versión microprogramada del procesador descrito en la figura precedente a tomar decisiones de tipo "salto condicional". El registro *F* contiene la dirección que en la memoria de microprograma ocupa la microinstrucción que en ese momento dispone del control de la senda de datos. La misión del sumador ahora añadido es sumar uno ("+ 1") a la dirección de la microinstrucción vigente, transmitiendo así la dirección de la inmediata microinstrucción secuencial del microprograma. Además, el sumador reinyecta en el sistema de control información tomada de la senda de datos. El signo del número alojado en el registro *C* ha de atravesar una puerta controlada por la línea *17* para acceder a la tercera entrada del sumador. Cuando esta línea se encuentra desactivada, se ejecuta la microinstrucción situada en la dirección secuencial sucesiva, $F + 1$ (pues $0 + F + 1$ es equivalente a $F + 1$). El sumador opera $0 + F + 1$ o $1 + F + 1$, según que el número alojado en el registro *C* sea, respectivamente, positivo o negativo. Así, con ayuda de la línea de control número *17* el microprogramador salta condicionalmente sobre una instrucción, dependiendo del signo del número situado en *C*. El microprograma de la operación de salto condicional se ve en la figura 9.

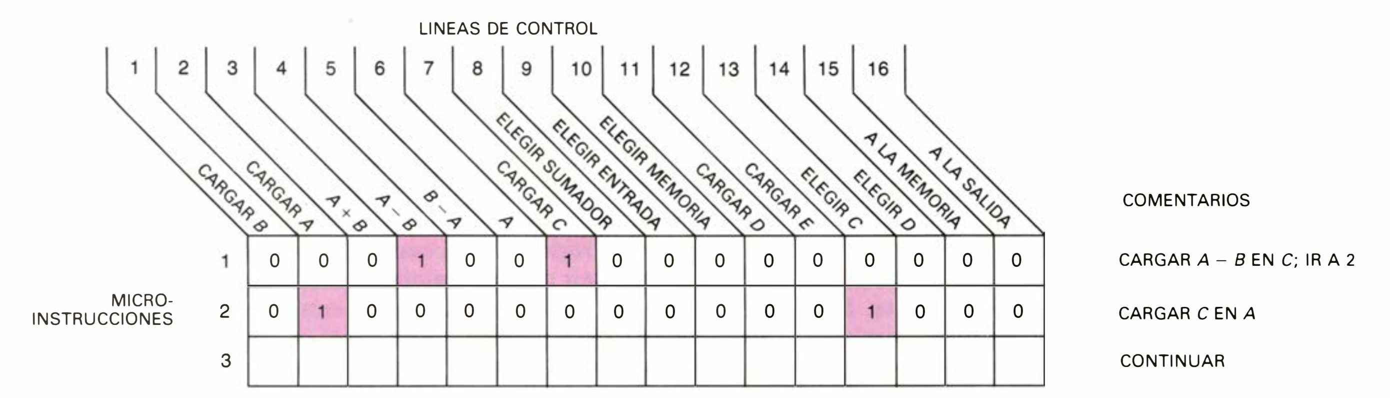
a que C sea mayor o igual que cero); se encuentra excitada (1) cuando el signo de C sea negativo (lo que significa que C es menor que cero).

La sustitución de los complejos circuitos de control por estructuras de memoria, de organización sencilla y repetitiva, conlleva dos ventajas fundamentales. Por una parte, se facilita la comprensión y construcción del sistema de control y, por otra, hace mucho más sencillo modificar el sistema mismo. Para enmendar un error en un sistema de control microprogramado, basta corregir los contenidos de las memorias. Supongamos, por ejemplo, que conviniera alterar el microprograma precedente, de modo que en lugar de cargar en E el contenido de D cuando A sea menor que B, la nueva versión haya de cargar en E el contenido de D en el caso de que B sea menor que A. La primera microinstrucción del microprograma primitivo carga en C el resultado de $A - B$. Ahora bien, si se desexcitase la línea de control 4 y se excitase la número 5, entonces la microinstrucción tendría como efecto cargar C con el resultado de $B - A$. En cualquiera de ambos casos, la microinstrucción siguiente determina si el número que hay en C es menor que cero, y salta si así es. Por tanto, podemos corregir la operación sin más que cambiar los estados de dos líneas en una microinstrucción. Con un sistema de control no microprogramado habría que reformar la estructura de los circuitos para efectuar la misma corrección.

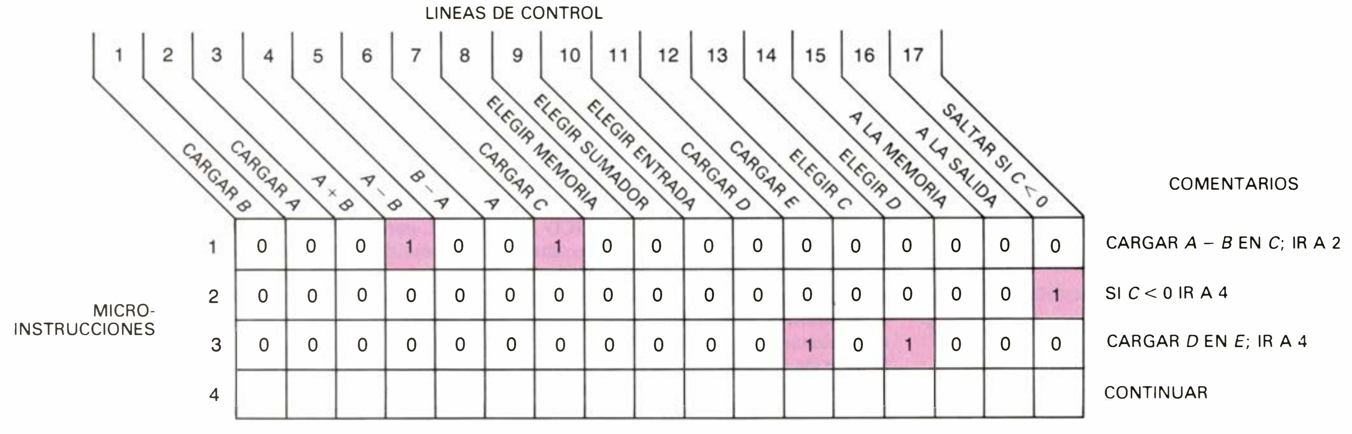
Además de reducir el costo de las correcciones a introducir en el sistema de control, la microprogramación permite cambiar por completo el sistema de control con sólo modificar el contenido total de la memoria asignada al microprograma. Puede así conseguirse que un mismo soporte físico sirva para muchas funciones diferentes.

Ya desde el principio saltaron a la vista las ventajas de la microprogramación. Empero, la idea de Wilkes era algo prematura, dado el nivel tecnológico del soporte físico de su tiempo. La microprogramación exige estructuras de memoria rápida y económica, de las que se carecía en el decenio de 1950 a 1960. Además, era convicción general en esa época que cualquier ordenador proyectado a fin de sacar partido de tan ingeniosa solución tendría forzosamente que ser más lento que las máquinas con sistema de control fundado en circuitos fijos. La microprogramación apenas si pasó de ser una curiosidad teórica durante más de diez años.

Hubo que esperar hasta los primeros años del decenio 1960-70 para que la microprogramación saltara a primer plano, gracias a una decisión de naturaleza comercial, y de trascendental importancia, adoptada por International Business Machines Corporation. En efecto, en los últimos años 50, la IBM constituía un aglomerado de pequeñas compañías, que comercializaban cada una su modelo de ordenador propio; ordenadores que no sólo diferían en tamaño, velocidad y precio, sino también en los sistemas de macroinstrucciones previstas para hacerlos funcionar. A causa de las diferencias en los sistemas de macroinstrucciones resultaba imposible hacer funcionar en unos tipos de ordenadores IBM programas preparados para otros tipos de la misma marca.



8. MICROPROGRAMA MODELO correspondiente a una sola macroinstrucción ejecutable por el procesador simplificado de la figura 6. El microprograma consiste en una sucesión de filas de una matriz de control, que es un casillero rectangular donde cada fila horizontal representa una microinstrucción y cada columna está asociada con una línea de control. En cada ciclo de reloj concreto, en las casillas correspondientes a líneas de control que deban encontrarse excitadas se anota un 1, mientras que si la línea ha de hallarse inerte se escribe un 0. (En la ilustración, los 1 van en fondo de color, y los 0, en fondo blanco.) La operación controlada por el microprograma consiste en restarle el número alojado en el registro *B* al situado en el acumulador *A* y almacenar después el resultado en *A*. Como sucedía con las macroinstrucciones, los comentarios de la derecha constituyen aquí meros recordatorios.

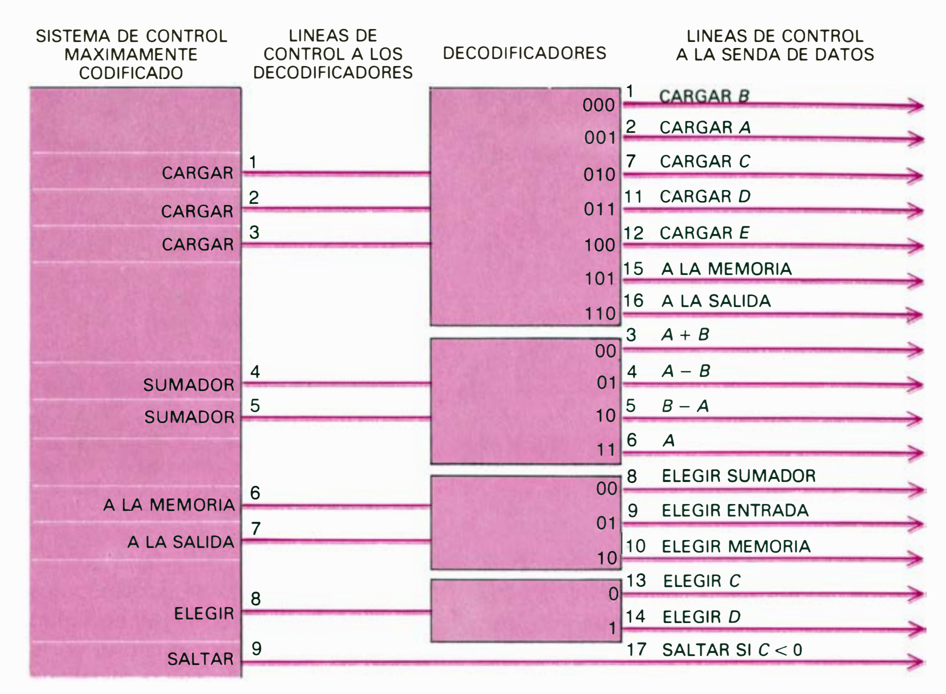


9. MICROPROGRAMA DE DECISION, que se vale del soporte físico ampliado de la figura 7. La operación controlada por el microprograma consiste en cargar en el registro *E* el número *D*, si es que *A* es mayor o igual que *B*. El salto condicional requiere una casilla adicional para dar la orden de saltar a la microinstrucción *4* si el número *C* es negativo (columna *17*). La macroinstrucción de salto condicional se reduce a optar entre la microinstrucción *3* y la *4*.

Evidentemente, resultaba mucho más sencillo vender ordenadores que operasen con un soporte lógico normalizado y se diferenciasen únicamente en tamaño, velocidad y precio. El problema consistía en convencer a los directivos de las compañías "subsidiarias" de que era posible para los diseñadores de tanto las máquinas más veloces como de las más económicas construir ordenadores eficientes fundados en sistemas de macroinstrucciones idénticas.

Se dio la circunstancia de que si los directivos de las subsidiarias alcanzaron a ponerse de acuerdo sobre la innovadora idea de familia de ordenadores con soporte lógico compatible fue tan sólo porque fueron persuadidos de que la microprogramación permitiría llevar a la práctica un proyecto semejante. Se argumentó que, siendo posible programar un mismo soporte físico para ejecutar sistemas de macroinstrucciones distintas, debería ser igualmente posible microprogramar ordenadores diferentes para ejecutar iguales sistemas de macroinstrucciones. La noción de familia compatible presentaba tan claras ventajas de índole comercial que IBM destinó grandes inversiones al desarrollo de tecnología de memorias rápidas y baratas para microprogramación. En 1964, anunciaba una familia de siete modelos de ordenador compatibles en soporte lógico, pero cuya velocidad oscilaba en un factor de 300 y, el precio, en un factor de 100.

No se decidió por sí misma, en IBM, que la microprogramación pudiera contribuir a resolver la difícil papeleta de hacer compatibles con las máquinas nuevas los programas disponibles para los viejos ordenadores. En consecuencia, varios modelos de la nueva familia IBM terminaron disponiendo de dos microprogramas: uno para el sistema de macroinstrucciones recién acuñado y otro para el antiguo sistema. Este proceder permitió que los clientes utilizasen sin modificación sus antiguos programas sobre el nuevo material. El método resultó tan bien acogido que durante los primeros años fue difícil decir si las nuevas máquinas pasaban más tiempo funcionando con el microprograma correspondiente a las nuevas macroinstrucciones o con el encargado de emular las antiguas.

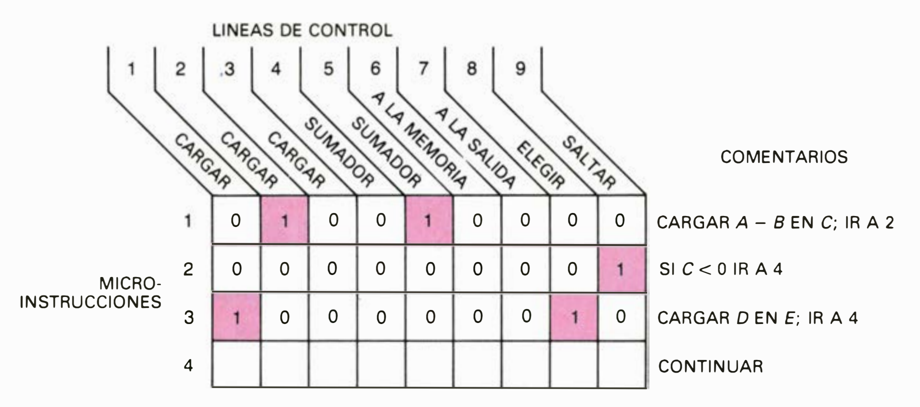


10. PARA REDUCIR EL NUMERO DE LINEAS DE CONTROL hay que codificar más las señales de control e introducir dispositivos apropiados para su decodificación. En general, el método consiste en buscar pares de líneas que muy raramente sean excitadas ambas en una misma microinstrucción, eliminar luego una de ellas y hacer que la otra cumpla una doble función. El soporte físico esquematizado en esta ilustración presenta una versión maximalmente codificada del procesador ya visto en las figuras 6 y 7. Las líneas de control que van a dar a la senda de datos han sido redispuestas, para subrayar que en el nuevo sistema de control las líneas se han codificado conjuntamente. A causa de esta mayor codificación, los sistemas de control son ahora, por lo general, más lentos que los minimalmente codificados (en los cuales, cada línea de control tiene asignada su propia casilla en el microprograma) porque el número de operaciones que pueden efectuarse simultáneamente es menor, y porque las señales de control han de pasar por los decodificadores. Además, los microprogramas resultantes tienden a constar de mayor número de microinstrucciones. Ahora bien, puesto que cada microinstrucción tiene un número menor de dígitos binarios, cabe reducir de forma importante los costos de memoria de microprograma.

La mayor parte de los principales fabricantes de ordenadores tomaron pronto la senda inaugurada por IBM, confiando en microcódigos para producir una familia de máquinas compatibles en soporte lógico que variaban ampliamente en capacidad, velocidad y precios. La principal dificultad con que tropezaron fue reproducir la onerosa tecnología inventada por los ingenieros de IBM; el problema terminó por quedar resuelto gracias a los rápidos progresos en tecnología microelectrónica. Durante más de un decenio, la industria electrónica ha duplicado la cantidad de información que se alberga en una sola pastilla de memoria. A lo largo del decenio de 1970 esta tecnología consolidó la posición de la microprogramación como método favorito de control. La excepción han sido los "superordenadores", que son los procesadores más veloces y caros. Unicamente en este caso particular sigue siendo preferido el sistema de control "a la medida" pues, aunque su construcción sea más difícil, consiente mayor velocidad.

Hace cosa de diez años, al hacer su aparición los primeros microprocesadores se planteó, como es natural, la cuestión de cómo diseñar sus sistemas de control. La historia se repitió y los primeros microprocesadores fueron construidos con sistemas de control no microprogramados. Sin embargo, conforme la tecnología ha ido madurando, las ventajas de la microprogramación han persuadido a casi todos los diseñadores de microordenadores a optar por la microprogramación.

En el ínterin, se ha desarrollado un tipo de microprogramación que ha ido alejándose de la idea de matriz simple, donde cada casilla está asociada con una línea de control individual. Las mejoras obtenidas son consecuencia, sobre todo, de los esfuerzos por reducir los costos de la memoria asignada al microprograma. Ocurre por lo general que, al ejecutar cualquier macroinstrucción individual, tan sólo están excitados unos cuantos componentes del soporte físico; mientras, las líneas de control conducentes a todos los demás elementos permanecen inactivas. Los microprogramas propenden, pues, a tener islas de unos salpicando un mar de ceros. Si los mismos objetivos pudieran conseguirse con menor masa de memoria (es decir, rebajando el número de ceros) el resultado sería una memoria más económica. La solución consiste en suministrar a la senda de datos la misma información codificando todavía más las señales de control. En general, el paso de codificación adicional estriba en buscar dos líneas de control que raramente se encuentran excitadas para una misma macroinstrucción, eliminar una de ellas y hacer que la otra cumpla dos funciones. Semejante proceder tiende a producir mayor número de mi-



11. EL MICROPROGRAMA MAXIMALMENTE CODIFICADO correspondiente al sistema de control de la figura 10 se dice que es "vertical" porque comúnmente consta de mayor número de instrucciones abreviadas, que lo hace más alto y estrecho que el correspondiente microprograma "horizontal". En este caso, la codificación pudo hacerse sin necesidad de recurrir a otras microinstrucciones adicionales.

croinstrucciones; a cambio, cada microinstrucción está formada por menos dígitos binarios. Un sistema de microinstrucciones al que se ha dado codificación maximal se denomina "vertical", porque los microprogramas resultantes acostumbran a ser altos y estrechos. Se dice, por el contrario, que un sistema de microinstrucciones minimalmente codificado es "horizontal", porque los microprogramas que producen son ordinariamente anchos y bajos. Las microinstrucciones minimalmente codificadas permiten, por lo general, mayor velocidad al ordenador, pues pueden ejecutarse simultáneamente mayor número de operaciones.

La creciente importancia de la microprogramación ha estimulado también el interés por dar con procedimientos que permitan a los usuarios del ordenador preparar sus propios microcódigos. Las ventajas que potencialmente presentan la transformación del ordenador, convirtiéndolo de instrumento de uso general en otro hecho a la medida para una determinada aplicación, ha determinado que la microprogramación por usuario sea una opción atractiva. Hasta ahora, los mejores resultados obtenidos por usuarios que han preparado sus propios microprogramas han alcanzado aumentos de velocidad que en algunos casos llegan a un factor de 10. En la práctica, la ganancia de velocidad depende de la senda de datos particular del ordenador microprogramado y de la aplicación de que se trate; en ciertos casos, las ganancias en velocidad son mucho más reducidas, si no llegan a perderse.

En la mayoría de los ordenadores microprogramados, las instrucciones se encuentran almacenadas en forma de memoria "de lectura solamente" (rom, o sea, read-only memory) cuyo contenido se determina en fábrica y permanece después inalterable. Otro tipo de ordenador microprogramable consiente modificar el contenido de la memoria de microprograma en tan sólo unos cuantos milisegundos, permaneciendo el ordenador en funcionamiento, entre tanto. De tales máquinas se dice que disponen de reserva de control inscriptible. La clave consiste en disponer en forma microprogramada de muchas aplicaciones, a punto ya para ser cargadas en cuanto sea necesario. Dotar al ordenador de una memoria de control inscriptible es uno de los procedimientos para suprimir la barrera que el soporte físico impone a la microprogramación por usuario.

A pesar de todos estos esfuerzos, la microprogramación ha seguido siendo tan difícil que pocos usuarios han sacado pleno partido de su potencial. De ahí que algunos investigadores se hayan aprestado a abordar el problema de simplificar la tarea de preparación de microcódigos. En muchos aspectos, la situación recuerda la que se planteó en los primeros días de la programación de ordenadores. El problema consistía entonces en la dificultad de estructurar miles de macroinstrucciones, articulándolas en un programa que funcionase sin tropiezos. El reto inspiró a los programadores la idea de escribir programas especiales que simplificasen la escritura de programas de aplicación. Estos programas especiales traducen lenguajes informáticos de alto nivel a macroinstrucciones más sencillas, que los ordenadores puedan ejecutar. Los lenguajes de alto nivel, como BASIC o Pascal, son más fáciles de comprender y utilizar, haciendo así menos onerosa la tarea de programación, si bien a costa de cierta pérdida de velocidad en la ejecución del programa.

Durante estos últimos años se ha estado trabajando en la creación de programas análogos, para traducción a microcódigo. El problema es en este caso más difícil, porque hay muchas más microinstrucciones potenciales que macroinstrucciones ordinarias. Un procesador microprogramado admite cómodamente 50 líneas de control; puesto que cada línea puede tener uno de dos estados (excitado o desexcitado), el número total de posibles microinstrucciones sería de 2^{50}, alrededor de 10^{15}. Además, como la velocidad del microprograma determina la velocidad del ordenador, resulta más importante todavía lograr en él un gran rendimiento. Se ha llegado así a la invención de lenguajes de microprogramación de alto nivel, tendentes a equilibrar facilidad de microprogramación con elevada velocidad y alto rendimiento. Los últimos desarrollos han sido estimulantes, pero la aceptación generalizada de la microprogramación entre los profesionales se encuentra todavía a varios años vista.

En el ínterin, los técnicos del Centro de Investigación Thomas J. Watson, mis colegas y yo, en el departamento de ingeniería eléctrica y ciencias de cómputo de la Universidad de California en Berkeley, hemos tomado un sendero diferente. Comenzamos desechando totalmente el sistema convencional de macroinstrucciones, y redactamos nuevos programas que traducen directamente a microcódigo los lenguajes normalizados de alto nivel. Las máquinas fundadas en este principio reciben la denominación genérica de RISC (reduced-instruction-set computers), es decir, ordenadores de sistemas de instrucciones reducidas. Para someter a prueba nuestras ideas, nuestro grupo de Berkeley ha diseñado su propio microprocesador, que llamamos RISC I. Hemos simplificado además la etapa de traducción, dando a las microinstrucciones codificación maximal, a fin de conseguir microprogramas del tipo más vertical posible. En efecto, todos los programas se convierten así en microprogramas.

Ensayos recientes efectuados por nuestro grupo y por los investigadores de IBM han demostrado que un ordenador provisto de sistema de instrucciones reducidas podría ser, en muchas aplicaciones, de tres a cinco veces más rápido que otro ordinario provisto de sistema de instrucciones complejas. En la Universidad de Stanford y en varios laboratorios industriales se ha acometido el desarrollo de proyectos para investigar este enfoque. Si los hallazgos provisionales llegan a substanciarse, podría llegar a emerger un tipo más sencillo de ordenador, que significaría un avance inesperado en la explotación del principio de microprogramación actualmente en marcha.

La escritura de la civilización del Valle del Indo

Una de las cuatro civilizaciones más antiguas del planeta posee una escritura que se ha resistido al desciframiento. La arqueología empieza a esbozar la posible solución

Walter A. Fairservis, Jr.

Las civilizaciones más antiguas son las de Sumer, Egipto, China y el Valle del Indo, en la región noroccidental del subcontinente ocupada en la actualidad por la India, Paquistán y Bangladesh. Los escritos de las tres primeras civilizaciones pueden leerse, pero las inscripciones de la cuarta siguen siendo enigmáticas en su mayor parte. En consecuencia, lo que se sabe de la civilización del Valle del Indo procede únicamente de los objetos materiales sacados a la luz por la arqueología. Es una ironía que, en una parte del mundo famosa por la antigüedad de su literatura, ni siquiera las fuentes más antiguas contengan alguna referencia válida a la primera gran cultura que se desarrolló allí.

La cultura de la civilización del Valle del Indo se denomina harappea, en razón de una de sus dos grandes ciudades: Mohenjo-Daro y Harappa. Las dificultades con las que se enfrentan quienes pretenden descifrar la escritura harappea parecen virtualmente insuperables a primera vista. Derivan en gran medida del uso limitado que hicieron los harappeos de su escritura. Sus "textos" consisten casi exclusivamente en breves inscripciones sobre sellos y un número igualmente limitado de grafitos sobre cerámica. Ninguna inscripción conocida consta de más de 21 signos y el texto medio alcanza cinco o seis solamente. Esto, unido al hecho de que muchos de los signos son pictográficos, ha favorecido que volara la imaginación de más de algún estudioso. Se han realizado denodados intentos por relacionar la escritura de Harappa con la minoica, la cananea, la hitita e incluso con la peculiar "escritura" de la isla de Pascua. Los estudios de los últimos años, fundamentados en lo que los datos arqueológicos revelan sobre la vida de Harappa, han facilitado un acercamiento más provechoso al problema del desciframiento. Mi intención es la de bosquejar aquí algunos de estos logros.

En las décadas que siguieron al descubrimiento de las dos grandes ciudades harappeas en los años 20 y, en particular, a partir de la segunda guerra mundial, los arqueólogos de la India y Paquistán han localizado unos mil yacimientos harappeos más. Estos se extienden en un arco abierto desde la India occidental, en los aledaños del río Narmada, hacia el norte, cruzando Gujarat y Kutch, a través de las regiones paquistaníes del Sind y el Punjab occidental, y hacia el Rajasthán indio y el Punjab oriental, hasta las puertas de Delhi. Se han encontrado otros asentamientos a lo largo de la costa del mar de Arabia hasta casi la frontera iraní, e incluso en el Beluchistán; una misión arqueológica francesa acaba de descubrir un asentamiento cerca del río Oxus, bien adentrado en Asia central.

La mayoría de los yacimientos harappeos son pequeños, con una superficie de una a dos hectáreas, y aledaños de ríos o arroyos. Hubo, al parecer, tres fases de asentamiento. Los yacimientos de la fase temprana se encuentran cerca o en los parajes fronterizos entre el subcontinente y el Irán. Los de la fase madura se hallan más desperdigados. La mayoría están en el propio valle del río Indo, pero yacimientos alejados como los del Beluchistán y el de las inmediaciones del Oxus pertenecen también a la fase madura. Los de la fase más tardía caen hacia el sur y el este del Indo. Los hallazgos en estos yacimientos más tardíos contienen también pruebas de que la cultura harappea más antigua se encontraba en un proceso de fusión con pastores y agricultores de la zona. Estos hallazgos sugieren que la civilización harappea no abocó a un final repentino, sino que desempeñó un papel en el desarrollo del estilo de vida aldeana que es tan característico del subcontinente en la actualidad.

¿Qué nos dice el registro arqueológico sobre la vida harappea? En primer lugar, nos indica que su pilar principal era la agricultura. Se cultivaban cereales y hortalizas, y la ganadería era fundamental para la economía agrícola. En segundo lugar, nos revela que la cultura material de los harappeos era sencilla, pero no carente de riquezas. Empleaban el cobre y el bronce para algunos utensilios y armas, y ocasionalmente trabajaban el oro y la plata en las cuentas que constituían su principal forma de joyería. Al mismo tiempo, la gran mayoría de sus artefactos constan de madera y hueso, concha, pedernal y arcilla, materiales todos que abundaban en la zona. Entre las cuentas, no obstante, había algunas talladas primorosamente en piedras semipreciosas más raras como el ágata,la cornalina y el lapislázuli. (El yacimiento del Oxus estaba también cerca de las primitivas minas de lapislázuli.)

Otro aspecto sorprendente de la vida harappea lo constituía el grado de normalización. En la arquitectura, los adobes tenían un formato normalizado y se disponían de un modo normalizado también. Los sistemas de drenaje y alcantarillado seguían modelos establecidos. Las viviendas eran de dimensiones normalizadas y las estructuras especiales (posiblemente de carácter público) estaban situadas con respecto a las privadas según planos establecidos. Otros aspectos del mismo fenómeno incluían los pesos y medidas normalizados, la cerámica, que era de forma y decoración normalizados, y artefactos normalizados como cazos, pesas de telar e incluso carretas de juguete. Al mismo tiempo, ciertos aspectos de la vida harappea anuncian la posterior cultura de la India, por ejemplo el uso de peina-

dos distintivos y de múltiples ajorcas y collares, e incluso el estilo de las figurinas harappeas.

Valga lo anterior por somero resumen de lo que se conoce. Lo que ignoramos puede ser de igual relevancia. No hay pruebas de rivalidad entre los diferentes estados harappeos, de guerras, de un comercio internacional importante o de los reyes y las cortes y los grandes complejos templarios tan característicos de las demás civilizaciones del mundo antiguo. La arqueología no nos revela casi nada de la religión harappea o de su organización social y política. Con todo, la civilización harappea fue importante para las culturas que la sucedieron. Los harappeos cosecharon el algodón y quizá el arroz, domesticaron la gallina y tal vez inventaron el juego del ajedrez y una de las dos grandes fuentes primitivas de energía no muscular: el molino de viento. (La otra la constituía el molino de agua.)

¿Cuándo floreció esta civilización enigmática? Aunque se discrepa en torno a esta cuestión, la fase harappea madura parece haberse extendido desde el 2200 al 1700 a. C. aproximadamente. La mayoría de los yacimientos harappeos parecen haber sido ocupados durante no más de 200 años; dan la impresión de un desarrollo de corta duración caracterizado por una organización sustancial de grupos y una interacción regular entre distintos asentamientos.

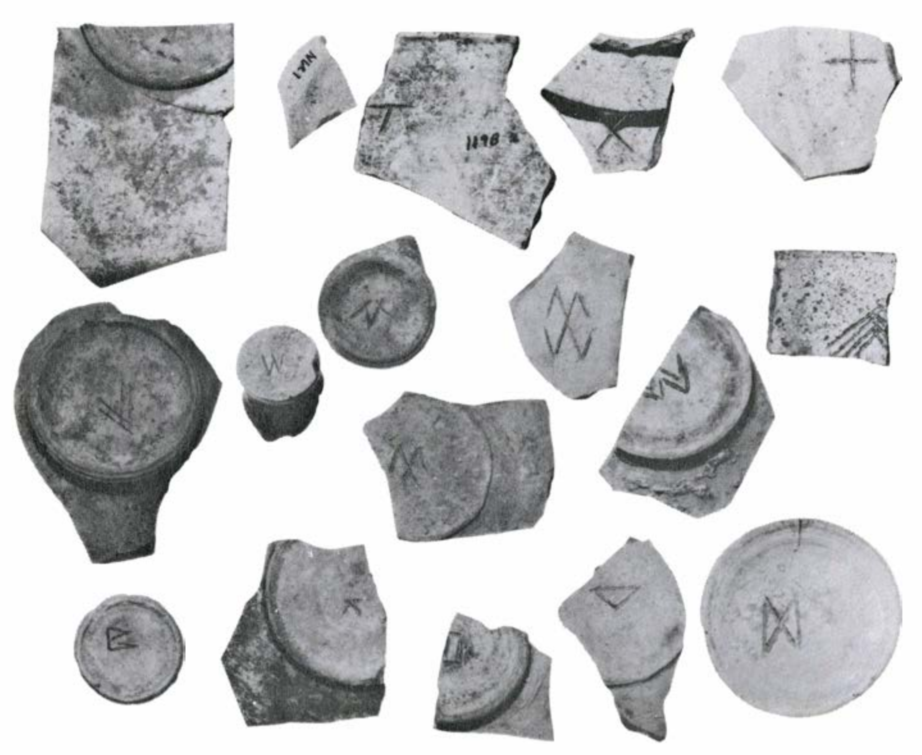
La labor de "leer" las inscripciones harappeas es lo suficientemente abstrusa como para hacer que muchos estudiosos la abandonen por imposible. La primera dificultad surge del hecho



1. SELLO DE LA CULTURA HARAPPEA de la colección arqueológica del Museo Nacional de Paquistán en Karachi. ("Harappea" viene de Harappa, ciudad del Valle del Indo.) El animal es un cebú. Los cuatro signos de la parte superior del sello se encuentran entre los que aparecen en el casillero de la figura 8. Son, de izquierda a derecha, el triple trazo del numeral 3, que tiene también el significado de "principal"; una figura de hoja del pipal y una flecha con el significado de "madre" o "señora"; una figura lineal con "cuernos", indicativa de un personaje de alcurnia y un recipiente con asas, sufijo honorífico con el significado de "elevado" o "superior". El sello pertenecía a un miembro destacado de la sociedad harappea, posiblemente una mujer.



2. "EL SEÑOR DE LAS BESTIAS", famoso sello de la colección arqueológica del Museo Nacional de la India en Nueva Delhi, muestra una figura humana sentada con las piernas cruzadas y con un tocado de búfalo de agua. Las "bestias" asociadas con esta figura son un rinoceronte y un búfalo de agua (*izquierda*), un elefante y un tigre (*derecha*) y una cabra con la cabeza vuelta (*abajo*). La cabra puede haber sido en principio una de una pareja, habiéndose perdido la segunda cuando se rompió el sello. La inscripción puede traducirse como "El Negro, el Búfalo Negro *an-i(l)*, el Altísimo, el Señor de los Caudillos".



3. GRAFITOS SOBRE CERAMICA. Los que aparecen en estos tejuelos hallados en el Beluchistán y otros constituyen la segunda fuente de signos harappeos. No son tan numerosos como las inscripciones sobre sellos e improntas, si bien nos indican que las inscripciones se leían de derecha a izquierda.

de que las inscripciones representan la escritura desconocida de una lengua desconocida para la que no hay textos bilingües, como la Piedra de Rosetta de Egipto o el monumento de Behistún del Irán. Una segunda dificultad reside en la ausencia de textos largos. Todo el corpus de la escritura harappea consta de unos 4000 sellos, improntas de sellos y grafitos sobre cerámica; además, muchos sellos están dañados y los grafitos cerámicos interrumpidos por fracturas. Una tercera dificultad surge del dato siguiente: la civilización harappea estaba alejada de las demás civilizaciones de su tiempo, no sólo desde un punto de vista geográfico, y se mantuvo aislada históricamente de los posteriores desarrollos culturales del subcontinente.

Afortunadamente los sellos harappeos están inscritos con escritura y representaciones pictóricas, y los motivos pictóricos pueden proporcionar las claves de lo que significa la escritura. La mayoría de los sellos harappeos son piezas cuadradas o rectangulares de esteatita, material que es fácil de grabar. En el reverso de la mayoría de los sellos hay una protuberancia taladrada con un agujero para un cordel de transporte; en el anverso hay una combinación de una representación y una inscripción, ambas en negativo. Los caracteres identificaban presumiblemente al propietario del sello: cuando se aplicaba el sello a la arcilla blanda, el objeto impreso se asociaba al dueño.

La mayoría de sellos harappeos muestran uno u otro de los dos motivos pictóricos distintivos. El primero acoge a un animal. Por lo general se trata de un toro cornilargo, pero también aparecen cebúes con su giba característica, búfalos de agua, cabras, toros cornicortos, rinocerontes, tigres, gaviales (los cocodrilos fluviales de la India) y elefantes. Cualquiera que sea el animal en cuestión, delante del mismo aparece un objeto; representado con animales salvajes o dañinos, es como una fuente. El objeto figurado con los animales domésticos es un cesto o (en especial ante el toro cornilargo) una vara larga, como la que vemos en procesiones.

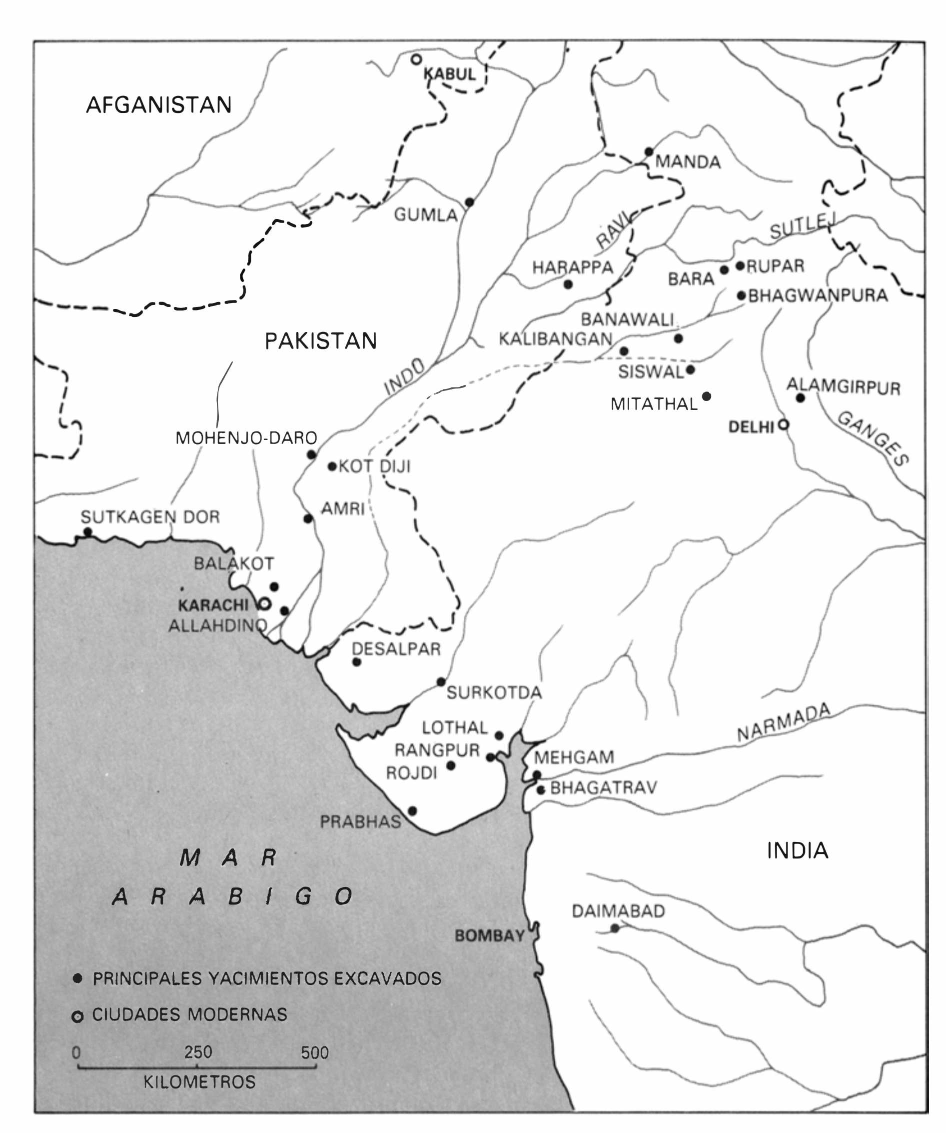
Algunos sellos de animales presentan sus temas en grupos. Un famoso ejemplo aparece centrado en torno a una figura antropomorfa, denominada a veces el "Señor de las Bestias", sentada con las piernas cruzadas y con un tocado de búfalo de agua. A la derecha de la figura se distinguen un elefante y un tigre; a la izquierda, un rinoceronte y un búfalo de agua. El sello está dañado, pero debajo y a la izquierda del pie de la figura central puede verse una cabra mirando hacia atrás y, a la derecha,

lo que podrían ser los cuernos de una figura caprina pareja, que por lo demás está borrada. Un sello menos elaborado se centra sobre un pipal o higuera sagrada (*Ficus religiosa*), de cuyo tronco salen las cabezas de dos toros cornilargos.

Otro "grupo animal" aparece representado en un sello prismático. (Algunos sellos harappeos están realizados en forma de prisma; los había también cilíndricos.) En una de sus tres caras están en fila un elefante, un rinoceronte y un tigre; una cuarta figura de la fila, demasiado erosionada, no permite su identificación. Probablemente se trata de la cara correspondiente al "animal salvaje" del sello. En la segunda cara hay otra procesión de cuatro animales, al parecer ninguno de ellos "salvaje". Sin embargo, en ambas caras aparece representado un quinto animal encima de la procesión faunística. Se trata de un gavial, que se autoevidencia como miembro del grupo "salvaje". En la cara de la composición salvaje, el gavial aparece figurado con un pez reconocible ante sus fauces y lo que puede ser otro pez tras él. La segunda cara no recoge pez reconocible alguno.

Un cuarto sello, aún más elaborado que el "Señor de las Bestias", introduce un elemento "cultual". Repite la representación de un árbol pipal, esta vez en su ángulo superior derecho. Entre las ramas del árbol se yergue una figura antropomórfica con cuernos. Frente a la figura astada hay otra de rodillas con faldas y, por tanto, probablemente femenina; a la izquierda de la figura arrodillada vemos una gran cabra. Siete figuras con faldas ocupan la mitad inferior del sello peinadas con una especie de larga "cola de caballo". Otra representación bastante horripilante, en un quinto sello, muestra varias figuras de similar tocado, una de ellas con faldas, que es atacada por un búfalo de agua.

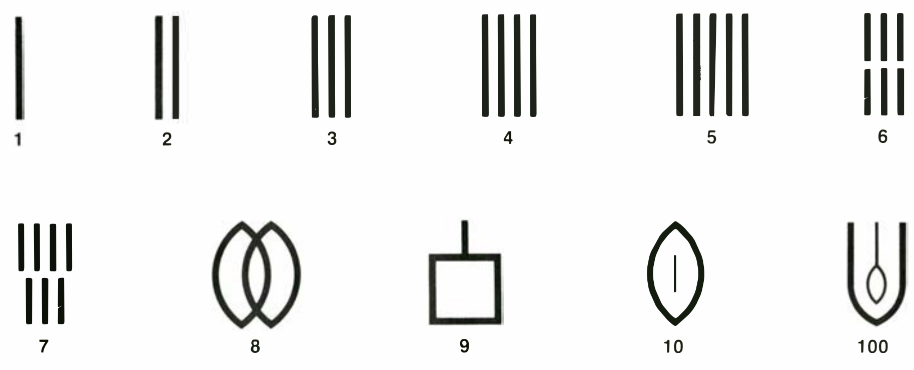
Estos motivos sugieren algo sobre la organización social harappea: que los individuos portadores de los sellos pertenecían a grupos que trascendían el marco familiar normal. Por ejemplo, todos los individuos con un rinoceronte en su sello podrían tener algún vínculo social en común. Lo que se compartía quizá fuera la pertenencia a algún grupo suprafamiliar: un clan o asociación. La existencia de escenas procesionales en determinados sellos, en los que parecen llevarse efigies animales a modo de estandartes, refuerza este concepto de los grupos suprafamiliares. Estos grupos a su vez pueden haber formado parte de un agrupamiento bipartito más amplio, como parece sugerir el motivo del "Señor de las Bestias", de un lado, y el motivo de los "adoradores de la deidad del Pipal", del otro. Esta estructura posible es conocida entre los antropólogos por "meidad": una sociedad caracterizada por la clasificación de los clanes u organizaciones subsidiarias similares en dos grupos o "mitades" que generalmente se enlazan entre sí.



4. SE HAN HALLADO YACIMIENTOS HARAPPEOS desde Bombay, el más meridional, hasta el Asia central, el más septentrional, y desde el Beluchistán, el más oriental, hasta Delhi, el más occidental. En este mapa se presentan veinticinco de los principales yacimientos excavados, incluidas las dos grandes ciudades de esta civilización: Mohenjo-Daro y Harappa. La mayoría de los asentamientos son pequeños.

Si se acepta la hipótesis de trabajo de que el material pictórico de cada sello identifica a su portador con respecto a su clan y meidad, se desprende lógicamente que la parte del sello dedicada a la inscripción harappea podría tener que ver con la identificación de su portador como individuo. La cartela podría dar, por ejemplo, el nombre del individuo, su ocupación, lugar de residencia, rango o título e información por el estilo. En apoyo de esta interpretación, el estudio de los textos de los sellos revela una considerable variación en la secuencia de los signos individuales y, sin embargo, una repetición frecuente de algunos de ellos. La hipótesis proporciona al presunto descifrador cierta base donde apoyarse. Del mismo modo que Michael Ventris sabía que al menos algunos de los textos en lineal B hallados en Creta eran inventarios de objetos materiales, así el descifrador de la escritura de Harappa puede suponer que lo que tiene que buscar son antropónimos y sus galas u otras precisiones que identifiquen al individuo.

Ya en la década de 1930, el estudioso británico G. R. Hunter había identificado un total de 396 signos harappeos independientes. Trabajos más recientes han añadido otros 23, de modo que sabemos que la escritura acoge un total de 419 signos. El análisis estadístico muestra que aparecen 13.376 veces en 2290 textos conocidos. De los 419, 113 signos se registran una sola vez, 47 se atestiguan por dos veces y 59 aparecen menos de cinco veces. En efecto,



5. LOS NUMERALES HARAPPEOS, parte en principio de un sistema octogesimal, constituyen una serie plausible de trazos verticales del uno al cinco y seis trazos breves para el seis. El signo para el siete es una serie similar de trazos menudos. Los signos para el ocho y el nueve son respectivamente los pictogramas de un sol doble y de un "poste de fundación". El signo para el diez es un solo sol con un breve trazo en su interior. Otro numeral identificado, el ciento, es un pictograma de un almirez con su mano.

esto significa que los 200 signos restantes tenían un uso más o menos general; el análisis muestra que la mitad completa de ellos son combinaciones de la otra mitad.

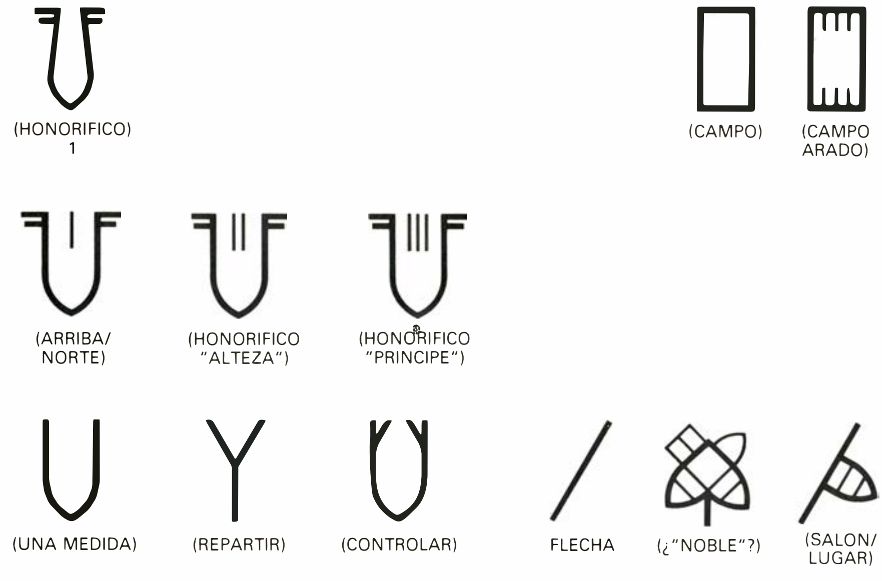
Estos hallazgos demuestran que la escritura harappea no era ni alfabética, como lo es el sánscrito, ni logográfica (esto es, con un carácter para cada palabra), como ocurre con el chino. Lo que sitúa a la escritura de Harappa en la categoría conocida como logosilábica, que quiere decir que algunos signos representan palabras y otros sirven puramente para sus valores silábicos o sonidos. Otros ejemplos de este tipo de escritura son los jeroglíficos egipcios, los ideógrafos sumerios primitivos y el japonés moderno. El hecho de que la mitad de los signos harappeos de uso habitual sean combinaciones de otros comúnmente utilizados, sugiere que los escritores explotaron dichos nexos para expresar ideas (tal como lo hace el chino cuando une los caracteres del sol y de la luna para representar la palabra "luminosidad") y para combinar sonidos silábicos al "deletrear" una palabra.

Una parte crucial de todo sistema gráfico la constituye la serie de mecanismos empleados para indicar el género, distinguir entre singular y plural, establecer la flexión de un verbo, etcétera. La identificación de estos recursos adelanta un largo camino hacia el establecimiento de las relaciones entre los grafemas, o componentes individuales, de las palabras, y la lengua que representan dichos grafemas. Ahora bien, los textos harappeos presentan ciertos signos emparejados regularmente, casi como la reiteración de títulos como "Su Majestad" y "Su Excelencia". Es asimismo notable que ciertos signos aparezcan en mitad de un texto, pero rara vez en su comienzo o su final; con otros, se da el fenómeno inverso.

El ordenamiento usual de los signos y la identidad de sus emparejamientos característicos pueden ser establecidos mediante un casillero a modo de rejilla. Dado que las inscripciones harappeas son por lo general escuetas, este casillero puede constar de un número relativamente limitado de columnas verticales. Se vio que bastaba con catorce columnas. De las primeras 17 inscripciones elegidas para su presentación horizontal en el casillero, la mayoría constaba de sólo cinco o seis signos; de este modo fueron asentados hacia el centro de la disposición en columnas, permitiendo que las líneas más largas se extendieran a la derecha o a la izquierda.

Al dar entrada en el casillero a los textos, quedó patente que determinados signos aparecían regularmente en la mayoría de ellos. Considérense los signos de "recipiente" de la columna 5, los signos de "la lazada" de la columna 8 y los signos del "doble trazo" de la columna 10. Cuando dos de estos tres signos aparecen en el mismo texto (como en las líneas *b, e, h* y *p*), siempre están en el mismo orden de derecha a izquierda entre sí, sin que importe qué otros signos estén incluidos en la inscripción. La regularidad de la posición regía evidentemente sus relaciones. Por consiguiente, al establecer el casillero de harappeo, estos signos se colocaron en las columnas señaladas arriba, aun cuando sólo uno de ellos apareciera en un texto determinado.

Lo que quedó claro después fue que dichos signos se contaban entre los más numerosos de cuantos componían los 17 textos seleccionados. El signo del recipiente de la columna 5 aparece allí 10 veces y tres variantes aparecen en otras dos columnas. El signo de la lazada de la columna 8 aparece seis veces, incluidas las variantes, y los dos breves trazos verticales de la columna 10, incluidas las variantes, aparecen ocho veces. Otros signos que utilizan los trazos verticales no están tan fijos en sus posiciones relativas. Mientras que el grupo de uno o dos trazos sueltos breves puede ser acomodado en la columna 10, los grupos de uno, dos, tres o más trazos largos caen fuera de las columnas centrales del casillero.



6. LOS SIGNOS COMBINADOS constituyen la mitad del total de los 200 signos de presencia más frecuente en la escritura harappea. Aquí se ofrecen cuatro ejemplos. El de la parte superior izquierda es uno de los signos más comunes, un pictograma de un recipiente con asas (identificado con el numeral arábigo 1). Debajo están tres signos combinados: el signo identificado con *l* más signos como los de los numerales harappeos 1, 2 y 3, pero de trazos breves. Debajo de los signos combinados se presentan, de izquierda a derecha, los pictogramas de un recipiente y de una horca, y de un tercero que combina los anteriores. En la parte superior derecha aparece el pictograma de un campo, un rectángulo sin más. A su lado hay un signo que combina el signo para campo con los trazos del numeral harappeo 6 para formar un nuevo pictograma con el sentido aparente de "campo labrado". En la parte inferior derecha, de izquierda a derecha, están los pictogramas para "flecha" y para "hoja del árbol pipal". El tercer signo, que combina "flecha" con el pictograma "hoja de pipal" parcial, parece significar "salón" o quizá "lugar".

A estas alturas, familiarizado ya con el aspecto de unos 50 signos harappeos, tal como se muestran en el casillero analítico, el lector puede preguntarse en qué dirección horizontal se han de "leer". De basarse sólo en las inscripciones de los sellos, ésta hubiera sido una cuestión difícil de contestar. Por fortuna, los grafitos inscritos sobre cerámica nos proporcionaron la respuesta. Los estudios de B. B. Lal e I. Mahadevan, del Servicio Arqueológico de la India, han mostrado que algunos grafitos presentan trazos que se solapan. Estos solapamientos demuestran que la dirección de la escritura era de derecha a izquierda. Por esto precisamente la columna 14 del casillero aparece en su margen izquierdo y la columna 1 en su margen derecho. La inversión coloca el orden de lectura en el modelo más familiar para nosotros de izquierda a derecha.

Pasamos ahora a la parte más complicada del problema del desciframiento: ¿Cuál era la lengua harappea? Cuando Ventris identificó la lengua de la Lineal B como griego, había superado su mayor dificultad. El registro arqueológico de la región ofrece un poco el mismo tipo de ayuda para quienes intentan averiguar cuál sería la lengua desconocida. Por ejemplo, la arqueología muestra que la civilización harappea no surgió de la noche a la mañana. Tuvo una larga progenie en las fronteras indoiranias, se difundió ampliamente y, con el tiempo, aportó su propia contribución al establecimiento de la India aldeana. Esta secuencia sugiere que la lengua hablada por los harappeos no puede haber desaparecido por completo del subcontinente.

De aceptarse esta presunción, ¿cuál de las tres principales familias lingüísticas habladas en la región podría relacionarse con el harappeo? Un candidato es el munda, una familia de lenguas (habladas en la India oriental) que parece emparentada con ciertas lenguas del Asia sudoriental. Sin embargo, los estudios sobre las formas más tempranas del munda apenas hallan en su vocabulario algo que se relacione con lo que la arqueología nos dice sobre la cultura de los harappeos. Otro candidato es el indoario, familia de lenguas que, según la tradición, entró en la India a mediados del segundo milenio a. C. Su expresión literaria más temprana, el Rig-Veda, describe no obstante una cultura básicamente centroasiática, bastante diferente de la harappea. El tercer candidato es el dravídico, lengua que se habla hoy, principalmente, en la India sudoriental y meridional, y que persiste en bolsas de la India septentrional y el Beluchistán.

Cabe, por supuesto, una cuarta posibilidad: que el harappeo no esté relacionado con ninguna de las citadas. Sin embargo, merece la pena considerar la candidatura del dravídico, mejor situado que el munda o el indoario. Además de las formas septentrionales (kurukh, malto y brahui), se hablan aún unas 25 lenguas dravídicas. Las familias principales (tamil, malayalam, kannada y telegu) son habladas por más de 100 millones de almas. El registro arqueológico, una vez más, nos libra de aceptar a ciegas la candidatura del dravídico.

En las excavaciones de 1974 en el yacimiento harappeo de Allahdino, en las cercanías de Karachi, se exhumó un fragmento de marfil con una sección semicircular. Había agujeros en una cara, taladrados, verosímilmente, para acoger clavijas; en el otro lado había dos ranuras paralelas a lo largo. El objeto era similar a un buen número de estacas o bastones de marfil excavados en la década de 1930 por uno de los últimos excavadores de la gran ciudad harappea de Mohenjo-Daro, E. J. H. Mackay. Los marfiles de Mohenjo-Daro no eran, como se ha asegurado, simples piezas de juegos. Por ejemplo, uno de éstos, de sección cuadrada, tenía tallados en una cara una serie de círculos y medias lunas alternados.

En alguno de los marfiles de Mohenjo-Daro hay un signo ya visto en el casillero (columna 6, línea *c*) y otro signo que se asemeja a cierta forma de planta. En distintos textos sigilares estos dos signos aparece asociados con trazos verticales, cuyo número va de uno a siete, y con otros cinco signos más. (Dos aparecen en el casillero en la columna 7, línea *c*, y columna 11, línea *q*.) Al igual que el segundo signo de los



7. PICTOGRAMAS ADICIONALES que pueden hallarse en los casilleros de la figura 8; incluyen los ilustrados aquí junto con algunas formas combinadas. En la fila superior, de izquierda a derecha, están "arco", "flecha", "hombre" y la combinación "arquero". En la segunda fila se exponen el "peine", una tercera figura humana lineal con un peine (combinación que parece indicar "escriba"), una cuarta figura lineal con "cuernos", una tercera figura lineal con un peinado elaborado ("mujer") y una figura lineal con una pértiga de hombro con carga (combinación que parece indicar "guardián"). En la tercera fila, de izquierda a derecha, se dibujan la "lanza" y la abreviatura para "lanza", "punta" (*color*), un título honorífico; una "lazada" con más brazos humanos, un segundo título honorífico referente al mando y los dos signos combinados. En la cuarta fila están dos tallos de cereal (ambos signos poseen el sentido intercambiable de "grano" o "mes"), un signo con el sentido de "río" o "agua" y una espiral con el sentido de "rodear" o "alrededores". En la fila inferior se recogen cinco pictogramas astronómicos. De izquierda a derecha van apareciendo, sucesivamente, el sol, la luna, el orto o el ocaso, cuarto creciente y una estrella.

marfiles de Mohenjo-Daro, el signo de la columna 6, línea *c*, parece representar una planta, probablemente la caña de un cereal. Considerando la asociación de ambos signos en las inscripciones sigilares con lo que parecen ser numerales y el grabado de uno de los marfiles con círculos y medias lunas, cuyo número se aproxima a un mes lunar (media luna) de 30 días solares (círculos), los dos signos vegetales podrían entenderse como la representación de una palabra que significase tanto grano como mes (o luna). Además, los trazos verticales asociados con ambos signos acaban tras alcanzar un total de siete, sugiriendo que algún otro signo servía para el numeral ocho y que posiblemente el sistema de numeración harappeo tenía la base ocho.

¿Qué lengua posee una palabra para grano que signifique también mes o luna y que esté asociada a un sistema numérico de base ocho? Para lo primero, *nel* significa arroz en cinco lenguas dravídicas y *nilā* o *nela* significa luna en tres de estas cinco lenguas y en otras cinco más también. Es también la palabra para mes en algunas de estas lenguas. En cuanto a lo segundo, un estudioso de la lengua dravídica, Kamir V. Zvelebil, ha señalado que el sistema numérico dravídico original era probablemente de base ocho: para contar hasta 10, en conformidad con el sistema actual de base decimal, se dice literalmente (en traducción) "uno", "dos", "tres", "cuatro", "cinco", "seis", "siete", "número", "muchos menos uno" y "muchos".

Durante años, los estudiosos han señalado en dravídico cierto número de homófonos: palabras con idéntico sonido, pero con significados distintos. Por ejemplo, la palabra para la pértiga de hombro de la que van suspendidos recipientes es *kā*. Es también ésta la palabra que significa guardián o protector. La palabra corriente para pez es *mīn*, que es también la palabra para estrella. La palabra para 100 es *nūṟu*, que es también la palabra para moler o pulverizar.

El principio homofóno o jeroglífico se da en cierto número de lenguas antiguas, incluidos los jeroglíficos de Egipto. El aspecto crucial del principio homofónico reside en que busca un equivalente silábico (digamos una representación de un pan para significar "pan") más que la pura representación de algo. Por ejemplo, un dinasta egipcio temprano, *Nr-mr*, era representado en jeroglíficos con el signo para un siluro (*nr*) y el signo para un cincel (*mr*).

		14	13	12	11
a	FEMD 606				
b	HAR 16				
c	HAR 99				
d	HAR 72				
e	FEMD 590				
f	MD-31 121				
g	MD-31 46				
h	MD-31 26				
i	HAR 110				
j	HAR 102				
k	FEMD 111				
l	MD-31 69				
m	HAR 69				
n	MD 650				
o	MD 405				
p	FEMD 48				
q	ME-31 110				

8. CASILLERO ANALITICO DE UN ANCHO DE 14 COLUMNAS. Contiene 17 textos sigilares, cuya dimensión va de dos a nueve signos. Cuando se asignaron los signos de "recipiente" a la columna 5, los signos de "lazada" y sus variantes a la columna 8 y los signos construidos con dos breves palotes verticales a la columna 10, quedó claro que, si dos o más de estos tres signos aparecían en una inscripción, siempre se sucedían en idéntico orden de derecha a izquierda entre sí, independientemente de los otros signos que aparecieran en el texto. Las líneas *b*, *e*, *h* y *p* nos ofrecen cuatro ejemplos. Los breves palotes verticales de

10	9	8	7	6	5	4	3	2	1

la columna 10 (líneas *b, d* y *j*) revelan una cualidad flexiva. El mismo signo aparece también en nexo en la columna 12, línea *d*, como adición al signo de "sol"; en la columna 11, línea *k*, como adición al signo "hombre" y en la columna 9, líneas *n* y *o*, como adiciones a los signos de "recipiente". El sentido del signo aislado y de sus combinaciones (líneas *i, n* y *o*) parece ser el de un posesivo; el sentido del signo de doble trazo parece ser el de un locativo. Dos de los textos, líneas *l* y *m*, no encajan dentro de la distribución central "normal" del casillero. Puede suponerse, pues, que no se trata de textos de formulación usual. Cuando se encuentran dos pictogramas de estos en una sola casilla, dicha práctica indica que los signos se hallan, por lo regular, emparejados en los textos sigilares. La lectura de cuatro de estos 17 textos (líneas *d, e, g* y *k*) se recogen en la figura 10. (La reconstrucción es de Andrew Christie.)

No se pretendía que la representación sugiriera que se llamaba "El Siluro-Cincel", sino más bien que su nombre sonaba como "siluro-cincel".

¿Sugiere esto alguna base para extraer un silabario harappeo? Voy a someterlo a comprobación en lo que sigue, avisando en primer lugar que la piedra de toque de todo desciframiento reside en su consistencia. Si se decide declarar a un determinado símbolo equivalente de un sonido o sentido elegidos, no puede cambiarse después para que represente algún otro sonido o sentido. La lógica interna es básica para todos los sistemas de escritura: cuando los descifradores alteran arbitrariamente los valores de los símbolos para adecuar el modelo que prefieren, su trabajo queda destruido.

Considérese como punto de partida el signo presentado en la columna 7, línea *h*. Pictográficamente podría suponerse que representa un almirez con su mano. Las reconstrucciones lingüísticas sugieren que a este utensilio se le daba en dravídico el valor silábico *nūṟu*. Como hemos visto, éste es también el sonido del verbo moler o pulverizar y el nombre del número "100". Considérese después el signo que figura de una forma tan prominente en la columna 5 del casillero. Hace algunos años Mahadevan conjeturó que este signo, que aparece con mayor frecuencia al final de los textos de los sellos, era una representación pictográfica de un recipiente con asas. Llegó a señalar que varias palabras dravídicas para este tipo de recipiente eran homófonas de palabras que significan macho, incluido el sufijo honorífico común *aṉ*, empleado con nombres personales masculinos al menos desde época tan temprana como la era cristiana primitiva. El hecho de que dicho signo clausure los textos sigilares que contienen, según la hipótesis, nombres propios, refuerza la conjetura de Mahadevan.

Ahora bien, otro signo clausurador del casillero (columna 5, línea *f*) es un pictograma de una figura humana con una pértiga de hombro con un recipiente colgante en cada extremo. Como hemos visto, la sílaba para una pértiga con recipientes sólo, *kā*, es homófona de las palabras dravídicas que significan "guardar" o "proteger". Además, una palabra dravídica común para hombre es *al*. La combinación de estos dos permitiría la reconstrucción del signo de la columna 5, línea *f*, como una palabra bisílaba, *kā-āl*. En las lenguas dravídicas, por motivos de eufonía, las vocales adyacentes de este tipo están separadas por una consonante,bien una *v*, bien una *y*. De este modo la palabra reconstruida, *kāvāl*, podría traducirse como "el que guarda o protege", especificación de identidad personal idónea para figurar en un sello privado.

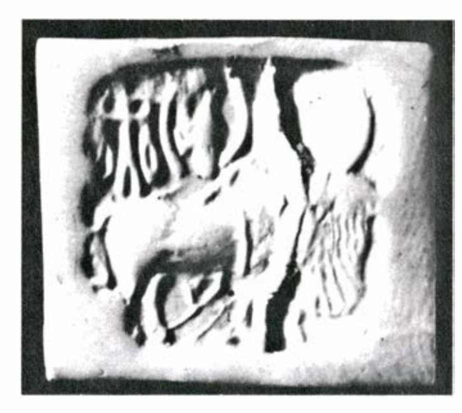
Para dar dos ejemplos más antes de volver a las lecturas comprobatorias de algunos textos sigilares del casillero, dos de los signos "numéricos" –el signo de la columna 12, línea *k*, que significa presumiblemente tres, y un segundo signo que consta de cuatro trazos verticales, con el significado posible de cuatro– presentan también homófonos plausibles. El primer homófono, *mu(n)*, se equipara con la palabra dravídica que significa "principal" (*muṉ*); el segundo homófono, *nāl*, se iguala con la palabra que significa "bueno" (*nal*), ambos adjetivos titulares lógicos.

Consideremos ahora dos textos sigilares breves y dos más largos del casillero. El texto designado con Mohenjo-Daro 3146 (línea *g*) consta de tres signos, quizás un instrumento musical, una lanza y un peine. Los valores silábicos dravídicos asignados son, en el mismo orden, *pan, ār* y *ki(r)*: en traducción la palabra "cantar", un honorífico de tercera persona y la palabra "marca", leído como *Panār-ki(r)*. La traducción propuesta es "sello de Panar" o literalmente "Marca del cantor". Después, el texto designado FEMD 590 (línea *e*) consta de cuatro signos. El primero, en forma de diamante, tiene el valor trisílabo *āra-man*, el segundo, dos breves líneas verticales, posee el valor *i(l)*, el tercero, una "lazada", presenta el valor trisilábico *piri(key)* y el cuarto es el mismo *ār* honorífico del texto precedente. Este texto se lee *Āra-man-i(l) pirikeyār*, con la traducción propuesta de "perteneciente a la noble casa (del) Pirikeyar" ("pirs" y "pirikeys" son caudillos).

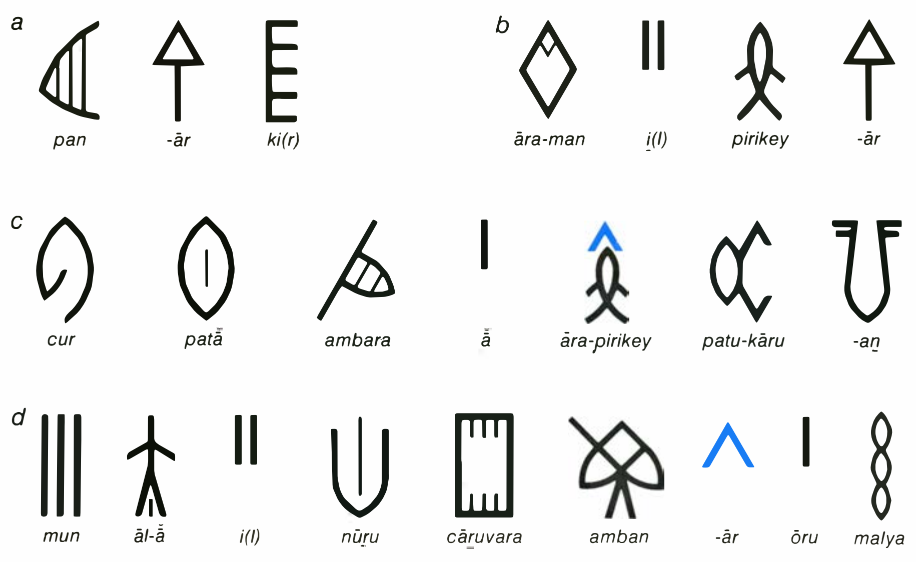
El tercer texto, Harappa 72 (línea *d*), consta de siete signos: una espiral inicial (*cur*), un óvalo que encierra un pequeño trazo vertical (*pată*), una hoja de pipal y una flecha en nexo (*ambara*), un breve trazo vertical aislado (*ă*), un lazo con una punta sobre el mismo (*āra-pirikey*), pinzas combinadas con el signo para el sol (*patu-kāru*) y, como signo de clausura, el recipiente con asas (*aṉ*). Es un trabalenguas: *Cur patămbara-ă āra-pirikey patu-kāran*. La traducción propuesta reza así: "Patukaran, caudillo (noble) poderoso de los asentamientos de los alrededores". El cuarto texto, FEMD 111 (línea *k*), es una de las más largas inscripciones sigilares traducidas, pues consta de nueve signos. En aras de la brevedad, sólo se ofrecerán las sílabas y su traducción propuesta: *Munālă-i(l) nūṟu cāṟuvara amban ārōru malya* o "perteneciente a Munala, señora de cien campos arados, noble primera dama".

Hasta el momento se han asignado valores silábicos a casi 100 signos en sellos y cerámica y para cada uno de ellos se han encontrado homófonos dravídicos apropiados. Se han propuesto traducciones como éstas para más de 100 textos sigilares. Estas van desde las sencillas declaraciones como "receptáculo de agua" (probablemente un nombre propio) hasta tan sonoras como "Arasamban, Alto Caudillo (de) Caudillos del Sudoeste, linaje de la Luna".

Lo que estas lecturas demuestran sobre los harappeos es que cierto número de individuos (Arasamban entre ellos) remontaban su linaje hasta figuras principales del cosmos como el Sol, la Luna y las Estrellas, y quizá también las Lluvias del Monzón. Estas pueden representar distinciones dentro de cada clan. Los caudillos aparecían asociados con *aramani* o moradas de caudillos, que pueden haber tenido tanto funciones residenciales como administrativas. Había además un lugar elevado (¿las "ciudadelas" identificadas por los arqueólogos?) que servían para una función especial, pero sin determinar hasta



9. HARAPPA Y MESOPOTAMIA. El contacto entre ambas queda probado por el descubrimiento de sellos harappeos en la región. El que se muestra arriba, con el sello a la izquierda y la impronta a la derecha, fue exhumado en 1975 en las ruinas de una casa del período cassita de Nippur, ciudad antigua del Iraq, por McGuire Gibson, del Instituto de Ciencias Orientales de la Universidad de Chicago.



10. CUATRO INSCRIPCIONES SIGILARES, junto con los valores silábicos de cada signo en dravídico, familia superviviente de las primitivas lenguas de la región. La inscripción más breve (*a*), en la lectura del autor, es "La marca del cantor". La inscripción que sigue (*b*), que repite el signo de la "lanza" honorífico de tercera persona, *ār*, se lee como "Perteneciente a la noble casa (del) Pirikeyar". La siguiente inscripción (*c*) tiene una longitud de siete signos e incluye una variante del signo de la "lazada" de la inscripción *b*. Reza "Patukaran, poderoso jefe (noble) de los asentamientos de los alrededores". La última inscripción (*d*) alcanza los nueve signos e incluye el signo de doble trazado *i(l)* de la inscripción *b* y una abreviatura de los signos de la "lanza" de *a* y *b* (*color*). Se lee así: "Perteneciente a Munala, señora de los cien campos arados, primera dama noble". Se han conseguido leer ya más de un centenar de textos sigilares. La mayoría de las lecturas se fundan en la hipótesis de que los signos de los sellos designan individuos.

el momento. Un área de reunión –un patio abierto o un salón hipóstilo (ambos conocidos por la arqueología)– sugiere que la asamblea de caudillos constituía una parte básica de la organización política harappea.

Entre otras clases de dirigentes conocidos a partir de los sellos estaban los responsables de asociaciones (¿gremios?) como las de los broncistas, los intendentes de los almacenes, los supervisores de las acequias y la de los propietarios de tierras, categoría que, según hemos visto, incluía mujeres. Las referencias religiosas son, sin embargo, escasas. Hay una posible "deidad con cuernos", referida como "el cobrizo" o "el rojo", y posiblemente una "deidad madre", pero por ahora estas distinciones no están confirmadas. Las figuras menores incluían tamborileros y cantantes; los tamborileros quizá para convocar asambleas y los cantantes para diversión o celebración de ceremonias.

Los sellos atestiguan que además había una clase de escribas, gente encargada de las pesas y medidas y supervisores de la distribución de los almacenes, molienda de la harina y probablemente cacerías. También había capitanes de embarcaciones y guardianes del fuego. Numerosos sellos presentan el signo *kā* que hace referencia a la guardia, no en sentido militar, sino más bien en términos de una responsabilidad para el cuidado de las cosechas y la conservación de las manadas y los rebaños. Una de las cuasipruebas gratificadoras de que los desciframientos van por buen camino procede del sello prismático donde se ve por dos veces un gavial encima de una procesión de animales del "clan". En dravídico la palabra para cocodrilo es *mutalai*. La palabra dravídica para caudillo principal, *mutali*, es un homófono cercano.

De este modo parece que la lengua de Mohenjo-Daro y Harappa hace unos 4000 años era una primitiva lengua dravídica, y que los escribas harappeos lucharon por poner esta lengua en forma gráfica como método para identificar la élite de la civilización del Valle del Indo. La civilización harappea, sin embargo, tenía una dispersión geográfica tal que un buen número de palabras no dravídicas deben haber entrado en la lengua, del mismo modo que palabras dravídicas fueron tomadas en préstamo por los hablantes del indoiranio más tarde.

Aún queda una gran parte por descifrar y sin duda se encontrarán lagunas en lo que se ha adelantado aquí. Lo que queda por hacer es aún más excitante, pues conduce a una meta que hasta hace bien poco parecía imposible de alcanzar: una información coherente sobre la política, organización social e ideología harappea y quizás una prueba aún más irrefutable de que, en tanto que emparentada con la India aldeana de hoy, la cultura harappea no desapareció jamás.

Ciencia y sociedad

Transformación eficaz

Uno de los problemas no resueltos que tiene planteada la biología es el de la conexión bioquímica entre el desarrollo de los seres vivos y la información genética contenida en su ADN. Se aborda introduciendo en las células germinales de un animal un gen implicado en el desarrollo que se haya aislado antes; a continuación, se observa qué efecto produce el gen en el animal. Para ello se examina la expresión del gen (proteínas) u otros parámetros biológicos del organismo alterados por la presencia de aquél. Lo ideal es que el gen introducido se convierta en parte permanente del genoma del animal, que pase a las generaciones siguientes y que se exprese en las mismas. La primera parte de este experimento no ofrece grandes dificultades; las técnicas de recombinación del ADN permiten aislar genes, modificarlos y prepararlos en grandes cantidades. La dificultad real se presenta al intentar introducir el gen así preparado en las células germinales, lograr su integración en el genoma sin ulteriores modificaciones y conseguir una integración estable para que lo hereden las generaciones sucesivas. Gerald M. Rubin y Allan C. Spradling, del Instituto Carnegie de Washington, acaban de resolver esta dificultad al desarrollar un método muy eficiente para introducir genes clonados en embriones de la mosca del vinagre *Drosophila melanogaster*. Según escriben en *Science* [*348*, 341-347 y 348-353 (1982)], han logrado introducir el gen *rosy* en embriones de moscas mutantes (con ojos castaños) y han obtenido adultos fértiles de cuya descendencia se han aislado linajes de moscas con los ojos rojos de las *D. melanogaster* normales.

D. melanogaster se cuenta entre los animales favoritos de los biólogos moleculares para el estudio del desarrollo, quienes han descubierto numerosos genes implicados en ese proceso. La simplicidad de la nueva técnica, que consiste en microinyectar, con finas agujas de vidrio estirado, pequeñas cantidades de solución de ADN en el polo germinal de los embriones de la mosca (una a dos horas tras la oviposición), explica que se esté utilizando en un centenar de laboratorios. Muchos investigadores habían dedicado en el pasado grandes esfuerzos a desarrollar un sistema de transformación en *Drosophila*, aunque sin éxito. El acierto de Rubin y Spradling estuvo en insertar antes el gen *rosy* en un nuevo vector natural: un transposón o elemento genético transponible. Un transposón es un fragmento de ADN que tiene la propiedad de moverse de un sitio a otro del genoma y puede insertarse él solo, y a veces conjuntamente con otros fragmentos de ADN incorporados dentro de él, en múltiples sitios distintos.

Rubin y Spradling trabajaron con la familia de transposones de *Drosophila* llamados elementos *P*. Rubin había identificado con anterioridad un elemento *P* particular, mucho más largo que los restantes y que parecía imprescindible para que tuviera lugar la transposición, tanto de él mismo como de los otros elementos *P* más cortos. Este elemento *P* es suficientemente largo para codificar una supuesta "transposasa", es decir, un enzima que catalizaría la transposición. Además, determinaría también una proteína represora que impediría la transposición. Los elementos *P* más cortos serían elementos "deficientes"; es decir: aunque capaces de transponer, carecerían de la información genética que determina estas proteínas. En opinión del grupo de Rubin, al introducir un elemento *P* entero en células germinales que no lo tienen se induce la síntesis de la transposasa; los elementos *P* presentes cambian de sitio con gran frecuencia. Al cabo de cierto tiempo, se acumula el represor y cesa la transposición.

En una primera serie de experimentos, Spradling y Rubin insertaron un elemento *P* completo en un plásmido bacteriano (una molécula circular de ADN capaz de replicarse en las bacterias). Obtuvieron cantidades grandes del plásmido recombinante, que inyectaron en embriones de *Drosophila*. Encontraron que el elemento *P* autocatalizaba su transposición al ADN genómico de las células germinales del embrión. Esta capacidad de autotransponerse sugería que podría también introducirse, en el genoma, un gen que interesara por alguna razón, con tal de que se hallara dentro de un elemento *P*, al que se utilizaría de vehículo para la transposición. Para conseguir una integración estable, se introduciría el gen en un elemento *P* deficiente, incapaz de sintetizar transposasa, y se añadiría al sistema una cantidad pequeña de la misma, suficiente para catalizar la transposición deseada, pero insuficiente para mantener la inestabilidad de los elementos *P* mucho tiempo.

De acuerdo con esta hipótesis, Rubin y Spradling prepararon un plásmido con un elemento *P* deficiente, en cuyo interior se había insertado un fragmento de ADN que contenía el gen *rosy*, clonado por W. Bender, de la Universidad de Harvard, y A. Chovnik, de la de Connecticut, y que determinaba el enzima xantín deshidrogenasa. Este enzima interviene en el metabolismo de las purinas; las moscas que carecen del mismo presentan ojos castaños. A continuación, inyectaron el plásmido (transposón *rosy*) en embriones de las moscas mutantes; para proporcionarles la transposasa necesaria recurrieron a dos procedimientos distintos. En el primero, inyectaron el plásmido en embriones mutantes derivados de un cruce entre machos que tenían en su genoma elementos *P* completos y hembras que no lo tenían. En el genoma de los híbridos, los elementos *P* transponen con gran frecuencia, pues en sus células hay transposasa proveniente de los elementos *P* paternos, que no se hallan "reprimidos" en las células del híbrido. El segundo procedimiento consistió en inyectar, en embriones mutantes, sin elementos *P* enteros o completos, una cantidad grande del transposón *rosy* y una cantidad pequeña de otro plásmido con el elemento *P* entero para que suministrara la transposasa necesaria. Habida cuenta del exceso de transposones *rosy* y la cantidad limitante de transposasa, se esperaba que muchas células incorporarían el transposón *rosy*, pero no el elemento *P* entero.

Los resultados fueron espectaculares. Ambos métodos funcionaron. Alrededor del 8 por ciento de los embriones inyectados se desarrollaron en adultos fértiles; de esta fracción, entre el 20 y el 50 por ciento dieron lugar a algún descendiente con los ojos de color rojo normal. Mas aún, la descendencia de estas moscas curadas de su defecto genético siguió manteniendo el color normal de ojos. Analizando el ADN de tales descendientes, se comprobó que el transposón *rosy* estaba integrado en el ADN cromosómico y era el responsable del cambio del color de

ojos. También se vio que, aunque los lugares de integración variaban de una estirpe transformada a otra, sólo una o a lo sumo dos copias del transposón se habían incorporado en los cromosomas. Finalmente, y como una ventaja adicional, también se observó que sólo el ADN del elemento *P* incompleto y de *rosy* estaban integrados; el ADN bacteriano del plásmido, que se encontraba alrededor del transposón, se había perdido. Para demostrar la validez general de su procedimiento, Rubin y Spradling llevaron a cabo la integración en los cromosomas de otro fragmento distinto de ADN de *Drosophila*.

Para los estudios del control de la expresión génica y para la genética molecular del desarrollo de *Drosophila*, disponer de un sistema de transformación sencillo y eficiente constituye una gran ventaja. Existe, además, la posibilidad de que los elementos *P* de *Drosophila*, o versiones modificadas de los mismos, funcionen en otros organismos. Y si resulta que los elementos *P* no pueden saltar las barreras interespecíficas, es muy probable que existan transposones análogos en otras especies. Rubin y Spradling piensan que la transferencia de genes mediante transposones permitirá modificar los genomas de animales y plantas; por ejemplo, para mejorar los rendimientos agrícolas o conferir resistencia a enfermedades. (Juan Modolell.)

W y Z

El descubrimiento de la partícula *W* supone un importante logro de la física experimental de partículas elementales. Se trata de una de las partículas conocidas por bosones vectoriales intermedios, o debilitones (en inglés, *weakons*). De acuerdo con la teoría dominante, que unifica el electromagnetismo y la fuerza nuclear débil, hay tres partículas de ese tipo: dos son portadoras de carga eléctrica (W^+ y W^-), mientras que la tercera es neutra (Z^0). En la teoría unificada electrodébil, los bosones actúan de intermediarios en las interacciones débiles entre partículas, al igual que el fotón transporta la fuerza electromagnética entre partículas dotadas de carga eléctrica. A diferencia del fotón, que carece de masa, los bosones *W* y *Z* tienen una masa enorme, más de 80 veces la del protón o del neutrón. De hecho esa masa ingente puso a los debilitones, durante mucho tiempo, fuera del alcance de los aceleradores de partículas [véase "Búsqueda de bosones vectoriales intermedios", por David B. Cline, Carlo Rubbia y Simon van der Meer; INVESTIGACIÓN Y CIENCIA, mayo de 1982].

El primer instrumento con energía suficiente para crear esas partículas se construyó hará un par de años en el CERN, Centro Europeo de Investigación Nuclear, con sede en Ginebra, modificando uno de sus aceleradores. La máquina, diseñada fundamentalmente para la búsqueda de bosones vectoriales intermedios, provoca la colisión entre haces de protones y antiprotones, que giran en sentidos opuestos, a una energía total de 540 GeV (miles de millones de electronvolt). Tras una serie de ensayos preliminares efectuados en 1981, a los que siguió un inesperado retraso debido a un accidente en uno de los dos principales detectores de partículas del acelerador, la búsqueda dio sus frutos a fines de 1982. A lo largo de un período de 30 días, entre noviembre y diciembre últimos, se registraron, para posterior análisis por ordenador, unos mil millones de colisiones protón-antiprotón a alta energía. En enero del año en curso, los físicos del CERN anunciaron que habían detectado pruebas de la existencia fugaz de partículas *W* en los rastros de una pequeña fracción de las colisiones.

Dos grupos experimentales llevaron la investigación. Dirigido uno por Carlo Rubbia, del CERN y de la Universidad de Harvard, lo formaban 126 físicos de 13 universidades y centros de investigación de Europa y los Estados Unidos. El otro, encabezado por Pierre Darriulat, lo componían 51 físicos de seis laboratorios europeos.

El primer grupo, que utilizó un gran detector de partículas multiuso, denominado UA-1 (por *U*nderground *A*rea One), ha identificado hasta ahora al menos cinco sucesos "que revelan la presunta rúbrica del bosón *W*". De acuerdo con el informe de los hallazgos de este grupo publicado por *Physics Letters*, se emplearon dos técnicas analíticas independientes para separar los sucesos de interés de entre los múltiples episodios. En la primera, el grupo seleccionó las huellas de un electrón, o positrón (la antipartícula del electrón), de alta energía, emitido de la zona de colisión a un ángulo muy pronunciado respecto del eje de los haces colisionantes. En la segunda, se buscaba pruebas de una falta significativa de energía, atribuible a la emisión de un neutrino desde la zona de choque. (Puesto que los neutrinos carecen de carga eléctrica, el sistema electromagnético de detección no los registra.)

En ambos casos, los experimentos hallaron cinco claros candidatos a *W* y un sexto probable. La comparación entre las dos series de datos reveló que uno y otro método de análisis habían identificado los mismos sucesos. Como señalan en su informe los experimentadores, "la presencia simultánea de un electrón y (un) neutrino de momentos aproximadamente iguales y opuestos en la dirección transversa... sugiere la presencia de una desintegración bicorpuscular". El fenómeno se interpreta como la desintegración espontánea de una partícula W^- de vida corta en un electrón y un antineutrino, o de una partícula W^+, igualmente fugaz, en un positrón y un neutrino. La masa de las partículas *W* que se deduce de las mediciones de los productos de desintegración es de 81 ± 5 GeV, lo que, según señala el grupo, está en "excelente concordancia" con la predicción de la teoría electrodébil.

El segundo grupo, que utiliza un detector más especializado, y en cierta medida más simple, de nombre UA-2, anunció en enero que había encontrado cuatro sucesos "que llevan la firma *W*". Pero puntualizaba: "se requieren más trabajos para confirmar este resultado provisional del UA-2". Ambos grupos tienen ahora la oportunidad de corroborar sus primeros hallazgos, puesto que el pasado mes de abril comenzó una serie de observaciones en el colisionador protón-antiprotón del CERN. Parte importante del programa es la búsqueda de Z^0, a la que se le supone una materialización 10 veces menos frecuente que la de las partículas *W*. La ventaja que presenta Z^0 tiene que ver con su desintegración en e^- y e^+, que, al poseer carga, su detección resulta más fácil.

Descubrimiento del fuego

En la reconstrucción humana de su propia historia, el dominio del fuego constituye un hito. Habilidad que comenzó, sin duda, con la captura y control del fuego espontáneo. Por los años 30 se descubrieron rastros de hogares y carbón vegetal, datados en unos 500.000 años, en el yacimiento de *Homo erectus* de Zhoukoudian, cerca de Beijing (Pekín). Luego se hallaron hogares en Vértesszöllös (Hungría) y Tautavel (Francia), lo que confirmó que la capacidad de controlar el fuego formaba parte del repertorio conductual de *H. erectus*.

Sin embargo, cuando se dieron los descubrimientos europeos, se le supo-

nía a *H. erectus* una antigüedad notablemente superior a los 500.000 años. En el lejano Oriente se habían hallado restos fechados en un millón de años y, en el este de Africa, se encontraron de hasta el millón y medio de antigüedad. Si la especie cumplía el papel cultural de un prometeo del Paleolítico Medio (el titán mítico, se recordará, robó el fuego pero no lo prendió), debía haber pruebas anteriores de que *H. erectus* sabía controlar el fuego.

Según parece, un equipo multidisciplinario de investigadores las ha recabado. Integran el grupo J. W. K. Harris, del Museo Nacional de Kenya, John Gowlett, del laboratorio de investigación arqueológica y artística de la Universidad de Oxford, Bernard Wood, especialista en anatomía de la Facultad de Medicina del Hospital Middlesex de Londres, y Derek Walton, del departamento de física de la Universidad McMaster. En su campaña por el Valle del Rift, en la región kenyata del Escarpe Laikipia, se concentraron en el lugar denominado Chesowanja, al este del lago Baringo, donde ya se había encontrado artefactos y restos animales (restos de homínidos incluidos). No tardaron mucho en desenterrar fragmentos de un material rojizo mineralizado. Walton, el geofísico, determinó que los fragmentos se componían de arcilla, cocida, hasta endurecerse, a una temperatura de 400 grados Celsius. La distribución de los pedazos cocidos indicaba una asociación directa con artefactos de piedra, guijarros sin pulimentar y huesos animales rotos. En un yacimiento, el material cocido guardaba gran semejanza con un hogar de construcción humana.

¿Acaso procedían los fragmentos de arcilla del incendio de la hierba, fenómeno habitual en la sabana de Africa oriental? No, pues en esos casos raramente se eleva la temperatura del suelo por encima de los 65 grados Celsius. ¿Se formaron por exposición a un calor más intenso, como la lenta consumición de tocones de árbol o la caída de un rayo? Quizá; no pueden descartarse sucesos aleatorios de ese tipo. Sin embargo, la asociación de los fragmentos de arcilla cocida con la construcción de hogares, la manufactura de artefactos y el consumo de alimentos apoya otra conclusión: la del control humano del fuego. ¿Cuánto hace de ello? Una estimación basada en la desintegración radiactiva del potasio en argón da la fecha de 1,4 millones de años atrás. ¿Quién encendió el fuego? Probablemente algún primitivo ejemplar de *H. erectus*.

ADN mitocondrial

El orgánulo celular que genera energía tiene su propio sistema genético. La organización de sus genes y su versión del código "universal" difieren de las del núcleo eucariota y las bacterias

Leslie A. Grivell

Hace unos 1500 o 2000 millones de años, cuando aún había poco oxígeno en la Tierra, una bacteria primitiva que malvivía de la fermentación anaeróbica de moléculas orgánicas absorbió una célula menor, que había adquirido la capacidad de respirar. El acontecimiento constituyó un punto crucial en la evolución orgánica. La respiración libera mucha más energía que la fermentación; la abundancia creciente de oxígeno en la atmósfera tuvo que ser la fuerza motriz impulsora de la relación simbiótica que se desarrolló entre las dos células, donde la aeróbica generaba energía en correspondencia al refugio y sustento ofrecidos por su anfitriona.

Con el tiempo, la célula absorbida, y otras como ella, se convirtieron en orgánulos subcelulares que las células anfitrionas transmitían a sus descendientes. Hubo un momento en que también cambiaron las células anfitrionas, desarrollando otras estructuras subcelulares y membranas internas, a la vez que repartían su material genético en cromosomas, encerrados en un núcleo. Estas células fueron las antecesoras de todas las modernas células eucariotas (nucleadas): protozoos, algas y las células de hongos, plantas y animales. Los descendientes actuales de aquellas bacterias simbióticas que respiraban son las mitocondrias, las centrales energéticas de la célula eucariota.

Las mitocondrias son orgánulos con forma ovalada, o agusanada, de medio micrometro de diámetro por dos o cinco de largo; este es, aproximadamente, el tamaño de muchas de las bacterias actuales. La mitocondria tiene una membrana externa y otra interna, extraordinariamente plegada, que encierra una matriz fluida. En el orgánulo se desarrolla la fosforilación oxidativa, fuente primaria de la energía celular. En la matriz, las moléculas orgánicas procedentes de la degradación de las materias alimenticias se oxidan en una serie de reacciones químicas conocidas por ciclo del ácido cítrico. Los electrones obtenidos durante la oxidación atraviesan una cadena de complejos enzimáticos respiratorios situados en la membrana interna, que dirigen la fosforilación del difosfato de adenosina en trifosfato de adenosina (ATP), el portador universal de energía en las células. El citoplasma de la célula eucariota (la región que queda fuera del núcleo) contiene un número variable de mitocondrias, de unas pocas a centenares; cuanto mayor es la demanda energética de la célula, mayor número de mitocondrias posee.

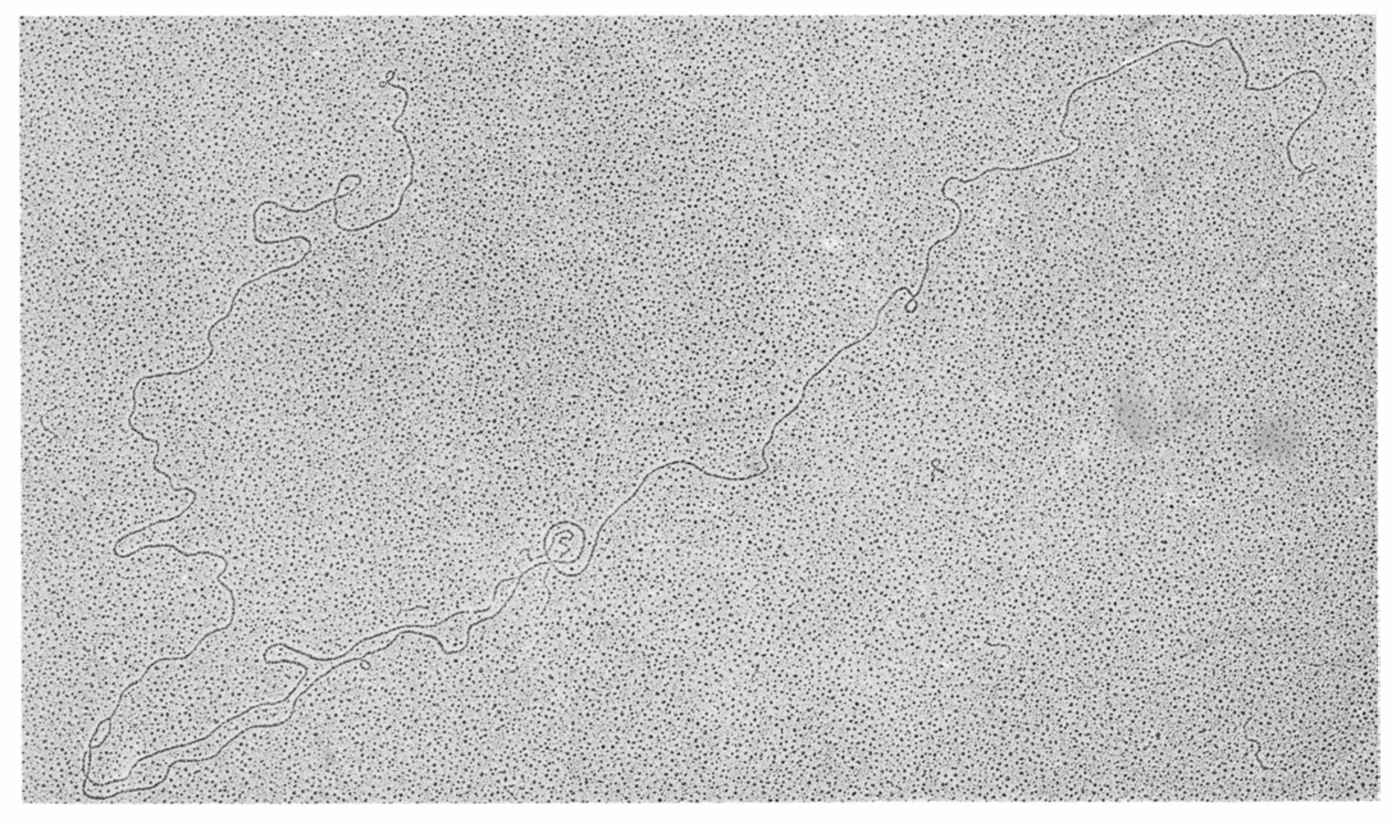
Las mitocondrias, quizá por ser descendientes de una bacteria "independiente", poseen su propio material genético y la maquinaria para expresarlo. La posibilidad de que los orgánulos pudieran tener sus propios genes, distintos de los nucleares, fue ya adelantada, en 1949, cuando Boris Ephrussi descubrió que la capacidad de la levadura de panadería para realizar la fosforilación oxidativa parecía estar controlada por algún factor citoplásmico, no nuclear. Sin embargo, hasta 1966 no se aisló y caracterizó el primer ADN mitocondrial de vertebrados (concretamente el del pollo). Desde entonces, en numerosos laboratorios de Europa y Estados Unidos se han venido estudiando intensamente los ADN mitocondriales de muchos organismos. Enseguida surgieron interrogantes a los que había que dar respuesta: ¿Para qué sirve el ADN mitocondrial? ¿Cómo se organiza y cómo se expresa? ¿Por qué la célula mantiene un segundo sistema genético, junto al nuclear? Ya es posible responder, en parte, a las dos primeras cuestiones y se puede especular, al menos, sobre la tercera.

En los dos últimos años se han adquirido nuevos e importantes datos, gracias, fundamentalmente, a la aplicación de la genética molecular junto con las nuevas técnicas de secuenciación rápida de nucleótidos (los "sillares" que componen el ADN). A principios de 1981, B. G. Barrell y sus colegas, del laboratorio Frederick Sanger, adscrito al Laboratorio de Biología Molecular del Consejo de Investigaciones Médicas de Cambridge, publicaron la secuencia nucleotídica completa del ADN mitocondrial humano; simultáneamente, Giuseppe Attardi, del Instituto de Tecnología de California, analizaba los productos de la transcripción del ADN mitocondrial humano. Poco después, el grupo de Barrell publicó la secuencia del ADN mitocondrial de bovinos y David A. Clayton y sus colegas, del Cal Tech, publicaron la de ratón. Mientras tanto, Alexander Tzagoloff y su grupo, de la Universidad de Columbia, desentrañaban la secuencia de una gran parte del genoma mitocondrial de levaduras. Parece confirmarse que los diversos ADN mitocondriales difieren de los de cualquier otro ADN conocido en varios aspectos fundamentales. Sus características, poco usuales, puede que sean las reminiscencias del primitivo organismo aeróbico que tiempo atrás colonizara las bacterias fermentadoras e iniciara, de esta forma, la "revolución eucariótica".

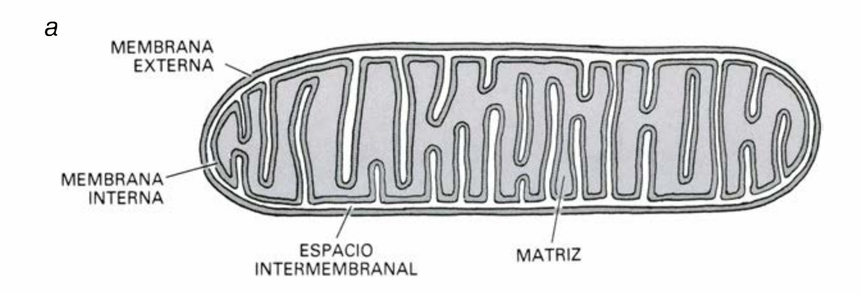
El genoma y sus proteínas

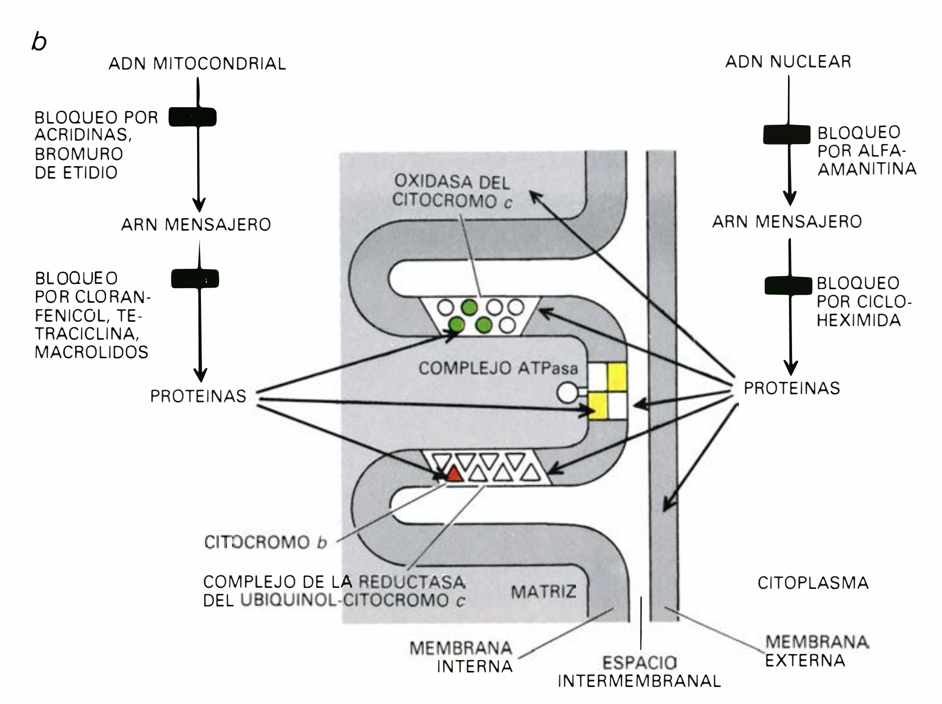
El ADN mitocondrial se asemeja al bacteriano en su calidad de doble hélice "desnuda"; el ADN nuclear eucariota está estrechamente asociado con proteínas y repartido entre varios cromosomas. Excepto en algunos protozoos, la hélice de ADN (de nuevo como en bacterias) forma un círculo. En la mayoría de las especies, sin embargo, el genoma mitocondrial es mucho menor que el de la bacteria más simple; su tamaño recuerda el genoma de muchos virus, aunque varía ampliamente según el organismo. La mayoría de las mitocondrias animales tienen sólo unos 15.000 nucleótidos; en levaduras, el

1. ADN MITOCONDRIAL de *Saccharomyces carlsbergensis*, la levadura de cerveza, aumentado unas 60.000 veces en una electromicrografía realizada por Ernst F. J. Van Bruggen, de la Universidad estatal de Groningen, Holanda. La molécula es un anillo "superenrollado" de ADN bicatenario, con una circunferencia de 26 micrometros. Consta de unos 75.000 nucleótidos (los "sillares" ADN, cuya secuencia determina la información genética).



2. GENOMA MITOCONDRIAL del protozoo *Tetrahymena pyriformis*. No se trata ahora de una molécula circular, sino lineal. El ADN aparece aumentado unas 26.000 veces en la micrografía tomada por Annika C. Arnberg y Van Bruggen. Se encuentran aquí en una etapa temprana de la replicación. Hacia la mitad de la molécula se ha formado un "ojo" replicativo, que se extiende a medida que el nuevo ADN se va sintetizando a lo largo de las cadenas.





3. ESTRUCTURA Y FUNCION de la mitocondria. Están determinadas conjuntamente por los genes del núcleo celular y de la mitocondria. El orgánulo posee una membrana externa y otra interna muy plegada que encierra una matriz fluida (*a*). En todos los organismos examinados hasta ahora, el ADN mitocondrial determina los ARN ribosómicos y de transferencia del orgánulo, elementos de la maquinaria de síntesis de proteínas. En las levaduras se conocen seis proteínas determinadas por el genoma mitocondrial (*b*): tres subunidades del enzima oxidasa del citocromo *c* (*verde*), dos subunidades de la ATPasa (*amarillo*) y el citocromo *b* (*rojo*). Las seis intervienen en la fosforilación oxidativa. La mayoría de las proteínas que constituyen el orgánulo, o que son necesarias para la expresión de su sistema genético, están determinadas por el núcleo. Si se bloquea con una droga selectiva el sistema de transcripción o de traducción puede saberse cuál de los dos sistemas, nuclear o mitocondrial, es responsable de una actividad sintética.

ADN mitocondrial es unas cinco veces más largo y, en los vegetales, se prolonga otras cinco veces más. Gran parte de la diferencia en longitud entre los ADN de levaduras y animal puede deberse a las notables diferencias existentes en la forma de organizar la misma serie básica de genes. Se ignora aún si esto mismo es cierto para el ADN mitocondrial de las plantas.

Aunque las mitocondrias tienen su propio material genético, no son genéticamente autosuficientes: la mayoría de las proteínas que se necesitan para la estructura y función del orgánulo están determinadas por el núcleo. El ADN mitocondrial posee información para la síntesis del ARN de sus propios ribosomas (estructuras donde se "fabrican" las proteínas) y para los ARN de transferencia (aquellos que transportan hasta los ribosomas los aminoácidos que van a formar las proteínas). En cuanto a proteínas, hasta el momento sólo se ha encontrado información para fabricar ARN mensajero correspondiente a algunas subunidades de los complejos enzimáticos respiratorios situados en la membrana interna del orgánulo. El resto de las proteínas de la membrana interna, las de la matriz (incluidos los enzimas del ciclo del ácido cítrico) y los de la membrana externa, están todas determinadas por genes nucleares; se sintetizan, por los ribosomas, en el citoplasma, y posteriormente se "exportan" a la mitocondria. El núcleo, además de suministrar a la mitocondria la mayoría de sus materiales estructurales, determina la síntesis de todos los enzimas necesarios para replicar el ADN y transcribirlo a ARN; además, el núcleo rige la síntesis de las proteínas que, junto con el ARN, forman los ribosomas, y codifica otros factores necesarios para la síntesis de las proteínas mitocondriales. En otras palabras, el sistema genético mitocondrial se mantiene sólo gracias a una considerable inversión por parte del núcleo y de la propia maquinaria celular de síntesis de proteínas.

Aunque con cierto margen de juego, la desigual contribución de los dos sistemas genéticos dentro de la célula está coordinada; de hecho, la célula puede pervivir algún tiempo aunque la actividad genética nuclear o mitocondrial esté bloqueada. La síntesis mitocondrial de proteínas se puede bloquear con antibióticos antibacterianos como el cloranfenicol, la tetraciclina y el grupo de los macrólidos; no se afecta con la cicloheximida, que inhibe la actividad de los ribosomas citoplásmicos. Puede diseñarse un experimento para marcar con isótopos radiactivos aquellas proteínas que se sintetizan mientras uno de los sistemas está bloqueado; de esta forma se logra determinar qué procesos corresponden a cada sistema.

Los experimentos de marcaje han permitido identificar los productos del ADN mitocondrial en varios organismos. Sin embargo, resultó más difícil identificar los genes responsables de los productos y cartografiarlos en el círculo de ADN, por lo menos mientras no se dispuso de los nuevos métodos de secuenciar ADN. Para identificar y cartografiar genes mediante análisis genético es necesario obtener y trabajar con mutaciones que interfieran la actividad de un gen, sin llegar a matar el organismo. En ausencia de oxígeno, una levadura puede llevar a cabo la fermentación, lo que significa sobrevivir sin mitocondrias funcionales. Ello permitió obtener y caracterizar mutaciones, en las levaduras, que en otros organismos serían letales. Ahí está la razón de que el ADN mitocondrial de levaduras fuera el primero en estudiarse a fondo.

La combinación de los métodos de análisis genético con los de secuenciación de ADN ha permitido ya realizar inventarios de los genes presentes en los genomas mitocondriales humano y de las levaduras, así como obtener mapas detallados de su distribución. Las mitocondrias de levaduras y del hombre tienen esencialmente los mismos genes, aunque se organizan de forma notablemente diferente.

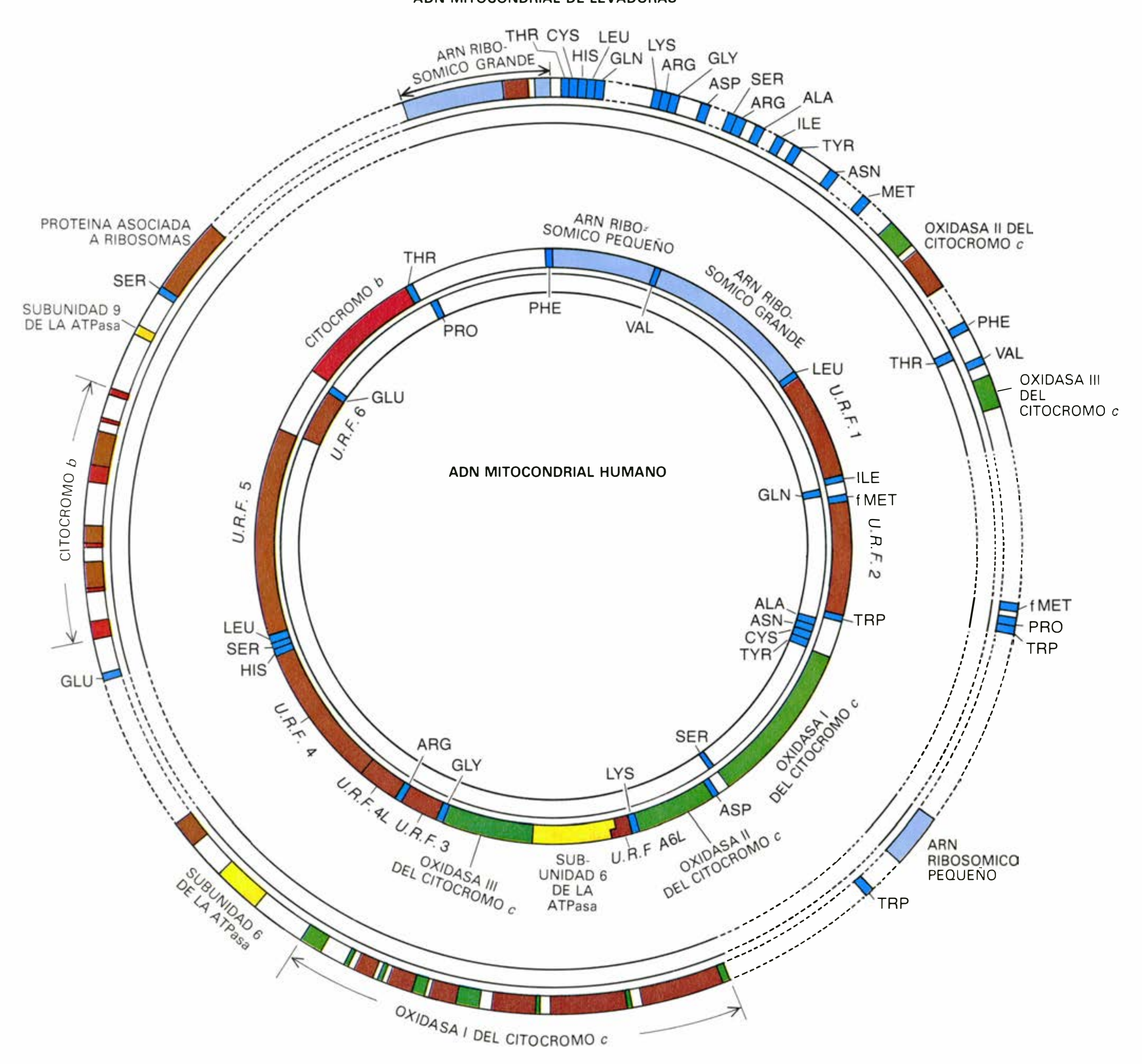
Organización del genoma

Los nucleótidos del ADN se caracterizan por la presencia de una base, de las cuatro posibles; las secuencias nucleotídicas de las dos cadenas se complementan: adenina (A) en una de las cadenas se empareja con timina (T) en la otra, y guanina (G) con citosina (C). Cada tres bases (triplete) de una de las

cadenas (la cadena determinante) forman una "palabra" clave (codón), que una vez transcrita a ARN, determina un aminoácido particular, o una señal para iniciar o terminar la traducción del mensaje. En la mayor parte del ADN cromosómico eucariota existen relativamente pocas palabras claves (codones). Las hileras de tripletes que constituyen un gen están flanqueadas por largos segmentos carentes de información y por regiones de señalización que no se transcriben a ARN, además de haber fragmentos que sí se transcriben a ARN mensajero pero no se traducen a proteína; a veces, los propios genes se encuentran fragmentados en trozos, separados por secuencias no informativas ("intrones").

El ADN mitocondrial humano es, en cambio, un modelo de economía. Tiene tan sólo 16.569 pares de bases, todos los cuales parecen tener alguna utilidad; ambas cadenas sirven, alternativamente, de cadena determinante. Además, los genes están tan "empaquetados" que entre ellos no hay prácticamente nucleótidos "inútiles". Hay genes para dos ARN ribosómicos, 22 ARN de transferencia y para cinco proteínas de membrana; hay, además, ocho posibles regiones informativas sin asignar. (Regiones donde se puede leer un texto coherente de forma continua, sin que esté interrumpido por codones de fin de mensaje. Estas regiones deben determinar alguna proteína que aún no se ha identificado.) Ninguno de estos genes contiene intrones. Hay muy poco ADN que determine ARN mensajero que no se traduzca a proteína.

Tanto en los genomas bacterianos



4. MITOCONDRIAS HUMANA Y DE LEVADURA; poseen, esencialmente, el mismo juego de genes, aunque difiera su organización. La ilustración esquematiza los dos ADN en forma de dos pares de anillos concéntricos (un anillo para cada cadena de la doble hélice), a escalas distintas: la molécula de ADN de levadura es, en realidad, unas cinco veces mayor que la humana. La secuencia completa de nucleótidos del genoma humano la dedujeron B. G. Barrell y sus colegas; también se conoce la mayor parte de la secuencia de la levadura *Saccharomyces cerevisiae*, aunque queden algunos fragmentos por secuenciar (*líneas discontinuas*). Los segmentos de color indican genes de proteínas conocidas (*verde, amarillo* y *rojo*) ARN ribosómico (*azul claro*) y ARN de transferencia (*azul oscuro*), cada uno de los cuales está señalado con la abreviatura de su aminoácido específico; los fragmentos marrones indican secuencias potencialmente informativas (U.R.F.), quizá genes de proteínas que aún no se han identificado. El genoma humano se muestra extraordinariamente compacto, con muy poco ADN no informativo entre los genes; el genoma de levaduras, por contra, ofrece largos fragmentos no informativos. Varios genes están interrumpidos por secuencias espaciadoras ("intrones"), algunas de las cuales incluyen secuencias potencialmente informativas.

→ OXIDASA I DEL CITOCROMO *c* →

TCATTCACTGATTTCCCCTATTCTCAGGCTACACCCTAGACCAAACCTACGCCAAAATCCATTTCACTATCATATTCATCGGCGTAAATC

7,100

TAACTTTCTTCCCACAACACTTTCTCGGCCTATCCGGAATGCCCCGACGTTACTCGGACTACCCCGATGCATACACCACATGAAACAT

7,200

CCTATCATCTGTAGGCTCATTCATTTCTCTAACAGCAGTAATATTAATAATTTTCATGATTTGAGAAGCCTTCGCTTCGAAGCGAAAAGT

7,300

CCTAATAGTAGAAGAACCCTCCATAAACCTGGAGTGACTATATGGATGCCCCCCACCCTACCACACATTCGAAGAACCCGTATACATA

7,400 ← ARN TRANSFERENTE (SER) ← →

AAATCTAGACAAAAAAGGAAGGAATCGAACCCCCCAAAGCTGGTTTCAAGCCAACCCCATGGCCTCCATGACTTTTTCAAAAAGGTA

→ ARN TRANSFERENTE (ASP) → OXIDASA II DEL CITOCROMO *c* →

TTAGAAAAACCATTTCATAACTTTGTCAAAGTTAAATTATAGGCTAAATCCTATATATCTTAATGGCACATGCAGCGCAAGTAGGTCTA

7,600

CAAGACGCTACTTCCCCTATCATAGAAGAGCTTATCACCTTTCATGATCACGCCCTCATAATCATTTTCCTTATCTGCTTCCTAGTCCTG

7,700

TATGCCCTTTTCCTAACACTCACAACAAAACTAACTAATACTAACATCTCAGACGCTCAGGAAATAGAAACCGTCTGAACTATCCTGCC

CGCCATCATCCTAGTCCTCATCGCCCTCCCATCCCTACGCATCCTTTACATAACAGACGAGGTCAACGATCCCTCCCTTACCATCAAA

7,800

TCAATTGGCCACCAATGGTACTGAACCTACGAGTACACCGACTACGGCGGACTAATCTTCAACTCCTACATACTTCCCCCATTATTCCT

7,900

AGAACCAGCCGACCTGCGACTCCTTGACGTTGACAATCGAGTAGTACTCCCGATTGAAGCCCCCATTCGTATAATAATTACATCACAA

8,000

GACGTCTTGCACTCATGAGCTGTCCCCACATTAGGCTTAAAAACAGATGCAATTCCCGGACGTCTAAACCAAACCACTTTCACCGCTA

8,100

CACGACCGGGGGTATACTACGGTCAATGCTCTGAAATCTGTGGAGCAAACCACAGTTTCATGCCCATCGTCCTAGAATTAATTCCCCT

8,200 ARN TRANSFERENTE (LYS) →

AAAAATCTTTGAAATAGGGCCCGTATTTACCCTATAGCACCCCCTCTACCCCCTCTAGAGCCCACTGTAAAGCTAACTTAGCATTAACC

U.R.F. A6L → 8,300

TTTTAAGTTAAAGATTAAGAGAACCAACACCTCTTTACAGTGAAATGCCCCAACTAAATACTACCGTATGGCCCACCATAATTACCCCC

8,400

ATACTCCTTACACTATTCCTCATCACCCAACTAAAAATATTAAACACAAACTACCACCTACCTCCCTCACCAAAGCCCATAAAAATAAA

SUBUNIDAD 6 DE LA ATPasa →

AAATTATAACAAACCCTGAGAACCAAAATGAACGAAAATCTGTTCGCTTCATTCATTGCCCCCACAATCCTAGGCCTACCCGCCGCAG

8,500

TACTGATCATTCTATTTCCCCCTCTATTGATCCCCACCTCCAAATATCTCATCAACAACCGACTAATCACCACCCAACAATGACTAATCA

8,600

AACTAACCTCAAAACAAATGATAACCATACACAACACTAAAGGACGAACCTGATCTCTTATACTAGTATCCTTAATCATTTTTATTGCC

8,700

ACAACTAACCTCCTCGGACTCCTGCCTCACTCATTTACACCAACCACCCAACTATCTATAAACCTAGCCATGGCCATCCCCTTATGAGC

8,800

GGGCACAGTGATTATAGGCTTTCGCTCTAAGATTAAAAATGCCCTAGCCCACTTCTTACCACAAGGCACACCTACACCCCTTATCCCC

8,900

ATACTAGTTATTATCGAAACCATCAGCCTACTCATTCAACCAATAGCCCTGGCCGTACGCCTAACCGCTAACATTACTGCAGGCCACC

9,000

TACTCATGCACCTAATTGGAAGCGCCACCCTAGCAATATCAACCATTAACCTTCCCTCTACACTTATCATCTTCACAATTCTAATTCTAC

9,100 →

TGACTATCCTAGAAATCGCTGTCGCCTTAATCCAAGCCTACGTTTTCACACTTCTAGTAAGCCTCTACCTGCACGACAACACATAATGA

→ OXIDASA III DEL CITOCROMO *c* → 9,200

CCCACCAATCACATGCCTATCATATAGTAAAACCCAGCCCATGACCCCTAACAGGGGCCCTCTCAGCCCTCCTAATGACCTCCGGCC

como en los eucarióticos no mitocondriales la transcripción de cada gen viene controlada, individualmente, por la interacción entre los enzimas de transcripción y secuencias señalizadoras presentes en las regiones no informativas del ADN. La escasez de ADN de este tipo en el genoma mitocondrial humano sugiere que la transcripción estará regulada de forma diferente. Gracias a los laboriosos trabajos de Attardi y sus colegas, del Instituto de Tecnología de California, comenzamos a vislumbrar cómo se desarrolla dicha regulación. Según parece hay un solo promotor, sitio donde se inicia la síntesis de ARN, en cada una de las cadenas del ADN mitocondrial. Por tanto, el producto inmediato de la transcripción es una copia completa de cada cadena en forma de ARN. Posteriormente, estos dos productos primarios se trocean, dando lugar a los ARN ribosómicos, de transferencia y mensajeros que, posteriormente, se traducirán a proteína; los cortes se producen a medida que se va sintetizando ARN. A falta de ADN no informativo, ¿cómo sabe el enzima dónde hay que cortar? Las secuencias correspondientes a los ARN de transferencia están íntimamente unidas a la de otros genes más largos; los cortes se producen de forma precisa al principio y al final de cada secuencia de ARN transferente. Ello da a entender que los productos primarios de la transcripción se plegarán de suerte tal que la propia estructura característica de los ARN de transferencia indicará los lugares que deben ser reconocidos por los enzimas encargados de procesar el ARN.

De la misma forma, la falta de espacios señalizadores cerca de los ARN mensajeros hace pensar en estrategias poco comunes para controlar la traducción. El codón para el aminoácido metionina sirve también de señal para iniciar la traducción. Los ARN mensajeros no mitocondriales van precedidos por secuencias que son reconocidas por los ribosomas, que se unen a ellas indicando así que el codón para metionina debe utilizarse como señal de iniciación. Los ribosomas mitocondriales humanos deben ser capaces de reconocer ciertos codones de metionina como codones iniciadores sin valerse de ninguna indicación precedente.

Parar la síntesis de proteína, al final del gen, constituye un problema más complejo. En los ARN mensajeros no mitocondriales, la terminación de la traducción viene indicada por alguno de los varios codones "terminadores", por ejemplo *UAA* (*U* simboliza el uracilo, análogo de la timina en el ARN). La mayoría de los mensajeros mitocondriales humanos o carecen de codón terminador (porque se solapan con el gen siguiente, generalmente alguno para un ARN transferente) o tienen uno que queda inutilizado al producirse el corte descrito anteriormente (mediado por el ARN de transferencia). Todos esos genes concluyen en *U* o en *UA*. ¿De qué forma, pues, se consigue un codón terminador? El ARN mensajero mitocondrial humano, como la mayoría de los mensajeros eucariotas, se "poliadenila", es decir, adquiere una larga cola de nucleótidos de adenina, inmediatamente después del corte. La cola de poli-*A* proporciona justamente lo que necesita la maquinaria de traducción: una señal de terminación *UAA*.

El ADN mitocondrial de las levaduras, al contrario que el humano, es un despilfarrador. Aunque el genoma de las levaduras no se ha secuenciado en su integridad, parece que hay información para dos ARN ribosómicos, unos 25 ARN de transferencia, seis proteínas de membrana y más de diez proteínas no identificadas. En las levaduras (más concretamente en *Saccharomyces cerevisiae*, la levadura de panadería), los genes están desperdigados sobre una cantidad de ADN cinco veces mayor. Largos segmentos de ADN rico en bases *A* y *T* se intercalan entre las secuencias informativas. Al menos tres genes están fragmentados por intrones y, además, al principio y al final de los genes se ubican largas secuencias que se transcriben pero no se traducen.

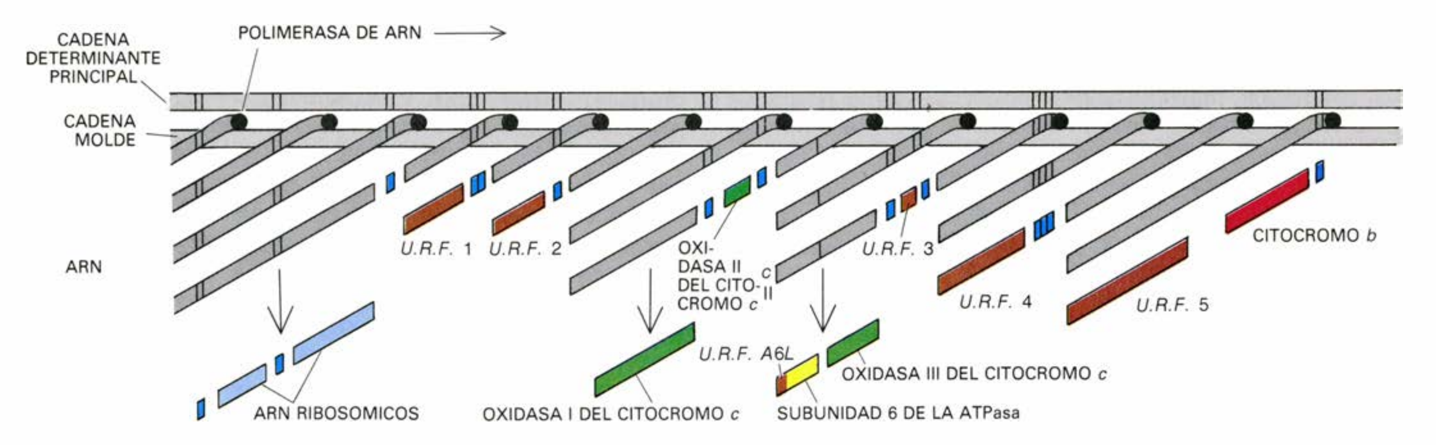
El código no es universal

Cada codón del ARN mensajero es "leído" por el anticodón de un ARN transferente, que es el que lleva cada aminoácido al lugar de la proteína que le corresponde. El código genético, que relaciona cada codón del ADN (o sus ARN mensajeros) con un aminoácido particular, es exactamente el mismo en virus, bacterias y animales, desde los gusanos hasta el hombre. Pero no es, como se creía, universal: las mitocondrias tienen su propia versión, que no es igual en todos los organismos. Las mitocondrias de los mamíferos leen *AGA* y *AGG* como señales terminadoras y no como el aminoácido arginina, y *AUA* lo traducen por metionina en vez de isoleucina; *AUA* y *AUU* sirven de señales de iniciación, en vez de *AUG*. Por otra parte, en las mitocondrias de los hongos *Neurospora* y *Aspergillus, AGA, AGG* y *AUA* se traducen de acuerdo con el código "universal". En las levaduras, como en mamíferos, *AUA* determina metionina, pero los cuatro codones que empiezan por *CU* se leen como treonina, en vez de leucina. En las mitocondrias de mamíferos, levaduras, *Neurospora* y *Aspergillus*, el aminoácido triptófano viene determinado por *UGA*, además de por el tradicional *UGG*, pero en el maíz el codón adicional para el triptófano parece ser *CGG*.

Aunque en total hay 61 codones (4^3, menos los tres codones terminadores), los organismos tienen bastantes menos ARN de transferencia, ya que un mismo ARN transferente puede leer más de un codón. Ello se debe a cierta libertad ("tambaleo") en el emparejamiento de la tercera base de los codones con determinadas bases situadas en la primera posición del anticodón. Según las reglas del tambaleo, propuestas hace varios años por F. H. C. Crick, 32 ARN de transferencia (pero no menos) estarían capacitados para leer 61 codones. No obstante, los análisis de secuencias demuestran que el ADN mitocondrial humano determina sólo 22 ARN de transferencia y, el ADN mitocondrial de las levaduras, 25. No parece que se utilicen ARN transferentes adicionales importados del citoplasma y, sin embargo, se leen 61 codones. ¿Cómo?

Según parece, cada uno de los 22 ARN transferentes de las mitocondrias de los mamíferos lee una "familia" de dos o cuatro codones sinónimos [*véase la figura 9*]. Los ARN de transferencia para familias de dos codones parecen seguir la regla convencional del tambaleo *G:U*, es decir, poseen anticodones, que empiezan por *G*, capaces de emparejarse con *U* y con *C*, o bien empiezan con *U*, que puede emparejarse con *G* y con *A*. Los ARN de transferencia pa-

5. FRAGMENTO DE LA SECUENCIA NUCLEOTIDICA del ADN mitocondrial humano, según datos tomados de Barrell y sus colegas. Se muestran sólo unos 2200 de los 16.569 nucleótidos, pero son suficientes para apreciar el grado de empaquetamiento de los genes. En una región (que empieza en el nucleótido 8527) hay dos genes que se solapan, es decir, los mismos nucleótidos desempeñan dos tareas distintas: determinan el final de la U.R.F. *A6L* y el comienzo del gen para la subunidad 6 de la ATPasa (como indican las dos llaves, las fases de lectura son diferentes.) La secuencia corresponde a la cadena utilizada como determinante por la mayoría de los genes indicados y tiene, pues, la misma secuencia que los correspondientes ARN mensajeros y de transferencia. El ARNt para el aminoácido serina está determinado de hecho por la cadena complementaria, que se transcribe en sentido opuesto, como señala su flecha.



6. ADN MITOCONDRIAL HUMANO transcribiéndose a ARN, que simultáneamente se procesa a ARN ribosómicos, de transferencia y mensajeros maduros, según Giuseppe Attardi, del Instituto de Tecnología de California. El diagrama ilustra cómo se transcribe la cadena determinante a un solo ARN, empezando en una región promotora única (*izquierda*). Este ARN se corta por los extremos de las secuencias de los ARN de transferencia (*rectángulos*), que aparecen en los flancos de casi todos los genes. El ARN inicial, producto directo de la transcripción, y los segmentos sucesivamente procesados se muestran en color gris. Los ARN maduros se indican de acuerdo con el código de colores ya especificado en las tres ilustraciones anteriores.

ra familias de cuatro codones, sin embargo, ignoran las reglas de Crick: por lo que se ve, leen los cuatro codones sinónimos. Sus anticodones tienen un *U* en la posición 5'. Ocurre entonces que *U* se empareja con cualquier base situada en el extremo 3' del codón, o bien sólo lee las dos primeras bases e ignora la última. Esta posibilidad, de leer dos de cada tres, fue sugerida hace varios años por Ulf Lagerkvist, de la Facultad de Medicina de la Universidad de Göteborg.

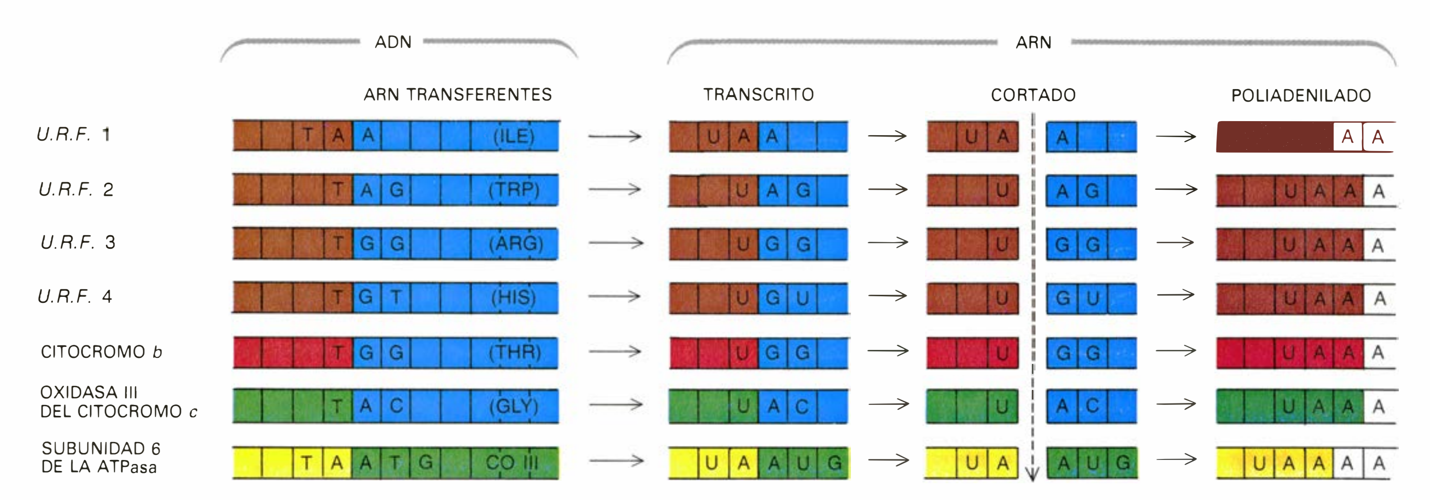
La cuestión que queda por resolver es cómo algunos ARN de transferencia con *U* en la posición "libre" ven limitada su capacidad de leer cuatro codones a dos. ¿Qué les impide reconocer codones que terminan en *U* o *C*, como los que terminan en *G* o *A*? En las mitocondrias de levadura y *Neurospora*, y en muchos ARN de transferencia no mitocondriales, el reconocimiento se impide modificando químicamente la *U* libre, pero hasta ahora los análisis de secuencias de los ARN de transferencia mitocondriales en animales no han revelado modificación alguna en dicha posición. No obstante, la mayoría de esos ARN de transferencia poseen una estructura poco convencional, y quizá la estructura influya en la especificidad del anticodón.

Intrones de las levaduras

Los genes fragmentados se descubrieron en 1977. Desde entonces hemos ido conociendo que muchos genes nucleares de mamíferos, aves y anfibios (y algunos de organismos eucariotas más sencillos) están fragmentados; secuencias no informativas ("intrones") se entremezclan con las que portan información para proteínas ("exones"). Cuando un gen de este tipo se expresa, primero se transcribe, intrones incluidos, a un ARN mensajero precursor. A continuación, este ARN se procesa (también se dice que madura): los intrones se eliminan y se empalman los exones, formándose un ARN mensajero maduro que se traduce a proteína. Los intrones de los genes cromosómicos, por lo que se sabe hasta ahora, no poseen información para fabricar nada.

En las mitocondrias de las levaduras las cosas son diferentes. Hace varios años, Piotr P. Slonimski y sus colaboradores, en el Centro de Genética Molecular de Gif-sur-Yvette, Francia, encontraron que determinadas mutaciones en uno de los intrones del gen para el citocromo *b* (un enzima respiratorio) de las mitocondrias de levaduras, afectaban al procesamiento del ARN producto de su transcripción. Tras elegantes experimentos genéticos, sumados al análisis de la secuencia del ADN, se llegó a la causa: la eliminación del segundo intrón del gen ¡dependía de una proteína codificada, en parte, por un segmento del mismo intrón!

Se han estudiado varias formas del gen del citocromo *b* en diferentes estirpes de levadura. La más compleja tiene seis exones (*B*1 a *B*6) y cinco intrones. Tres de los intrones poseen largas re-



7. FALTAN CODONES DE PARO, que señalan la finalización de la traducción a proteínas, en siete ARN mitocondriales humanos. En algunos, la secuencia "entra" en la secuencia informativa situada a continuación; en otros, el codón de paro se corta, y por tanto se deshace, durante la maduración del ARN. La necesaria señal de paro se adquiere de forma poco convencional, por poliadenilación, fenómeno normal del procesamiento del ARN por el que, para formar un ARN mensajero, a cada ARN se le añade una cola de nucleótidos de adenina (*A*). Con ello, las *A* completan el codón de paro *UAA*.

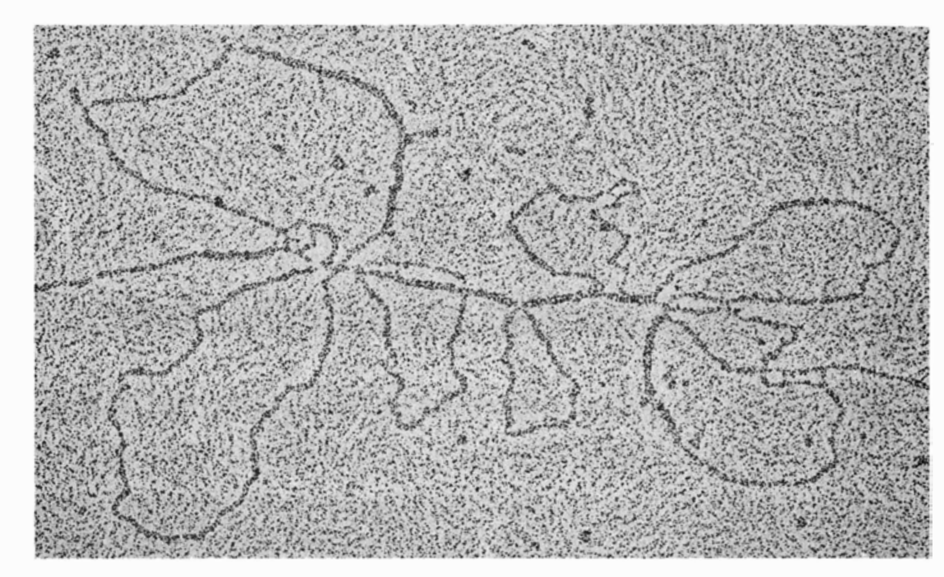
giones potencialmente informativas. El gen completo se transcribe a un largo producto primario. A continuación, se elimina el intrón número 1. Este paso conlleva la unión de los exones *B*1 y *B*2 y, paralelamente, el que los tripletes del intrón número 2 adquieran la misma fase de lectura que los exones que le preceden.

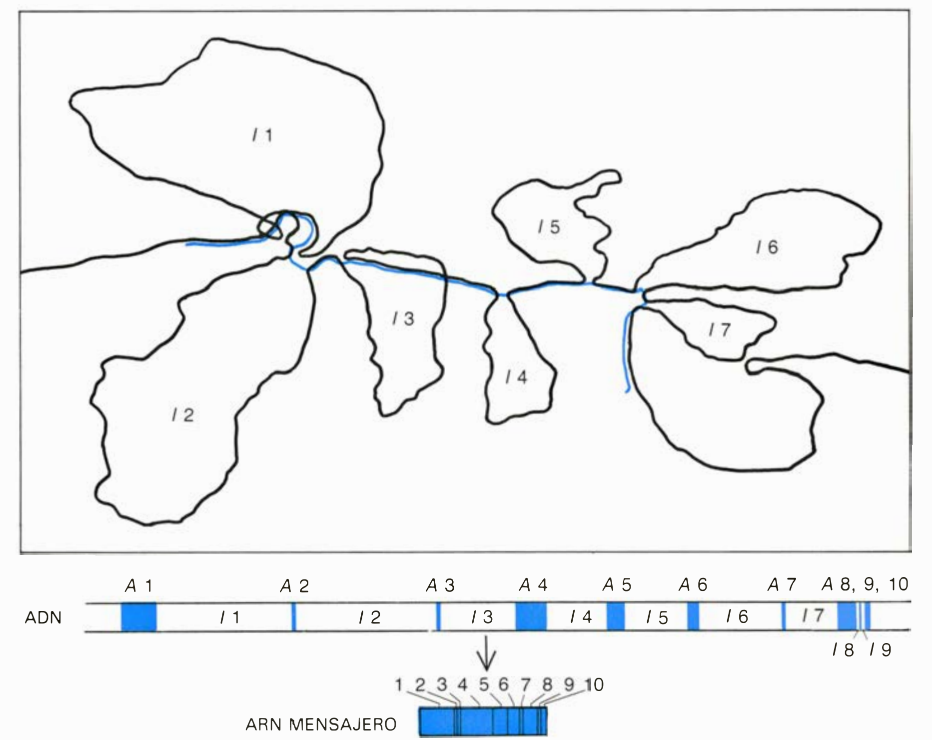
Ahora, los exones *B*1, *B*2 y el fragmento del intrón número 2 en fase hacen las veces de un ARN mensajero. Se traduce a una proteína cuyo extremo amino comprende los primeros 143 aminoácidos del citocromo *b* (determinados por *B*1 y *B*2) y cuyo extremo carboxilo viene determinado por el texto del intrón. Esta singular proteína ayuda a catalizar la eliminación del intrón número 2. En el proceso destruye su propio ARN mensajero, limitando por tanto su propio nivel en la mitocondria. Slonimski llamó a este fenómeno "homeostasis de procesamiento", y "madurasa" a la proteína determinada por la combinación exón-intrón, indicando que podría tratarse, no de un enzima capaz por sí mismo de catalizar el procesamiento, sino más bien de un factor que modificase en alguna forma la especificidad de un enzima procesador preexistente, determinado probablemente por un gen nuclear.

Pequeños anillos de ARN

La presencia de una madurasa del ARN explica claramente el comportamiento de ciertas mitocondrias mutantes, con lesiones en el intrón número 2. Tales mutantes no pueden procesar los productos del gen del citocromo *b*, sino que acumulan largos ARN de los cuales sólo se elimina el intrón número 1. Estos mutantes generan grandes cantidades de proteínas que reaccionan con anticuerpos contra el citocromo *b*, pero no son citocromo *b*. La síntesis de estas proteínas se bloquea si en el exón *B*1 se introducen mutaciones que provoquen el final de la síntesis de proteínas. Tales proteínas deben ser, por tanto, madurasas defectuosas procedentes de la traducción de los exones *B*1 y *B*2 y del mencionado fragmento del intrón número 2; su exceso de producción es el resultado de su incapacidad para destruir el ARN mensajero que dirige su síntesis.

La homeostasis de procesamiento no está, probablemente, limitada al segundo intrón del gen del citocromo *b*. Los intrones números 3 y 4 de este gen y los primeros cuatro intrones del gen para la subunidad I de la oxidasa del citocromo *c* (otra proteína de la membrana interna) también presentan largas regiones potencialmente informativas, y hay pruebas de que en su procesamiento pueden estar involucradas madurasas específicas de intrones.





8. GEN FRAGMENTADO del ADN mitocondrial de levaduras, cartografiado por microscopía electrónica. La micrografía, realizada por Arnberg y Van Bruggen (de la Universidad de Groningen), y la interpretación del mapa, en la parte inferior, muestran una molécula híbrida: una molécula de ADN monocatenario que contiene el gen de la subunidad I de la oxidasa del citocromo *c* (*línea negra del diagrama*) que se ha dejado hibridar con el ARN mensajero (*color*) transcrito a partir de dicho ADN. En el ADN, las secuencias que determinan proteínas están interrumpidas por intrones. Las secuencias del ADN monocatenario que se complementan en el ARN se emparejan con éstas y forman regiones bicatenarias; los intrones, que no están representados en el ARN, forman bucles monocatenarios. Cuando un gen como éste se expresa, el ARN primario "madura", esto es, se eliminan los intrones y quedan únicamente los "exones", o regiones con información para fabricar proteínas. La información recabada de los experimentos de hibridación, junto con los datos de secuencias obtenidos por Alexander Tzagoloff, de la Universidad de Columbia, y por Lambert A. M. Hensgens, en el laboratorio del autor, han permitido deducir la organización de los exones (de *A*1 a *A*10), de los intrones (de *I*1 a *I*9) y del ARN mensajero maduro.

Ciertos procesos de maduración generan, en las mitocondrias, pequeños anillos atípicos de ARN. Annika C. Arnberg, de la Universidad estatal de Groningen, los descubrió cuando colaboraba con Gertjan van Ommen y el autor en estudios de electromicroscopía para cartografiar los ARN mitocondriales. Los anillos resultaron ser intrones escindidos de los ARN precursores y circularizados, al parecer, mediante enlaces químicos covalentes. Se han encontrado anillos correspondientes a tres intrones del gen para la subunidad I de la oxidasa del citocromo *c* y a un intrón (el ya mencionado nú-

mero 2) del citocromo *b*. Aparentemente, el anillo del citocromo *b* y otro perteneciente a un intrón con una región potencialmente informativa se escinden con el concurso de madurasas determinadas por las mitocondrias, mientras que la escisión de los otros dos puede depender de enzimas procesadores determinados por el núcleo.

¿Por qué intrones?

Hace tiempo, Crick lanzó la idea de que durante el procesamiento podrían originarse anillos de ARN si un enzima procesador pudiese pegar los dos extremos de un intrón de la misma forma que une dos exones. Hasta el momento sólo se ha descrito otro caso donde el procesamiento origine ARN circulares. Hace varios años, Thomas R. Cech y sus colegas, de la Universidad de Colorado en Boulder, observaron que en el protozoo *Tetrahymena*, durante el procesamiento de uno de los ARN ribosómicos, uno de los intrones escindidos formaba un anillo. Cuando el grupo de Cech intentó identificar el enzima presuntamente implicado en el proceso de escisión y circularización, llegaron a la sorprendente conclusión de que no había tal enzima: la eliminación y circularización parecían depender de propiedades intrínsecas del mismo ARN. El ARN de *Tetrahymena* es el único ejemplo conocido de lo que Cech llama un ribozima: un ARN que se autoprocesa. Un conocimiento más profundo de los anillos mitocondriales que no requieren madurasas para la escisión debe revelar si en esos casos también están implicados los ribozimas.

La madurasa asociada con el intrón número 2 se distingue por una propiedad peculiar: no es indispensable. Abundan las estirpes de levaduras con versiones más cortas del gen para el citocromo *b*; a menudo, estos genes carecen de los tres primeros intrones, de forma que los exones *B*1 a *B*4 se han fusionado. La información contenida originalmente en los intrones que faltan no se encuentra en ninguna otra parte del genoma mitocondrial de las levaduras, lo cual sugiere que la madurasa es el equivalente mitocondrial del peor tipo de burócrata: aquél que existe sólo para regular su propia existencia. En realidad, los tres intrones ausentes son miembros de una clase de intrones mitocondriales de levaduras definidos como "opcionales" por Johan Sanders cuando, hace pocos años, trabajaba en nuestro laboratorio de la Universidad de Amsterdam. Existen otros ejemplos en los genes para el mayor de los dos ARN ribosómicos y para la subunidad I de la oxidasa del citocromo *c*. Para clasificar un intrón como opcional basta encontrar una estirpe que carezca de él y sea funcional. Hasta ahora sólo se ha examinado un número limitado de estirpes de levaduras; quizá se descubra que todos los intrones mitocondriales de levaduras son opcionales. La naturaleza opcional de unos intrones, aunque sean pocos, plantea nuevas perspectivas sobre algunas cuestiones básicas. ¿Por qué hay genes fragmentados y otros que no lo son? ¿Por qué el mismo gen existe tanto en la forma fragmentada como en la continua? ¿De dónde proceden los intrones opcionales, y, en general, los intrones?

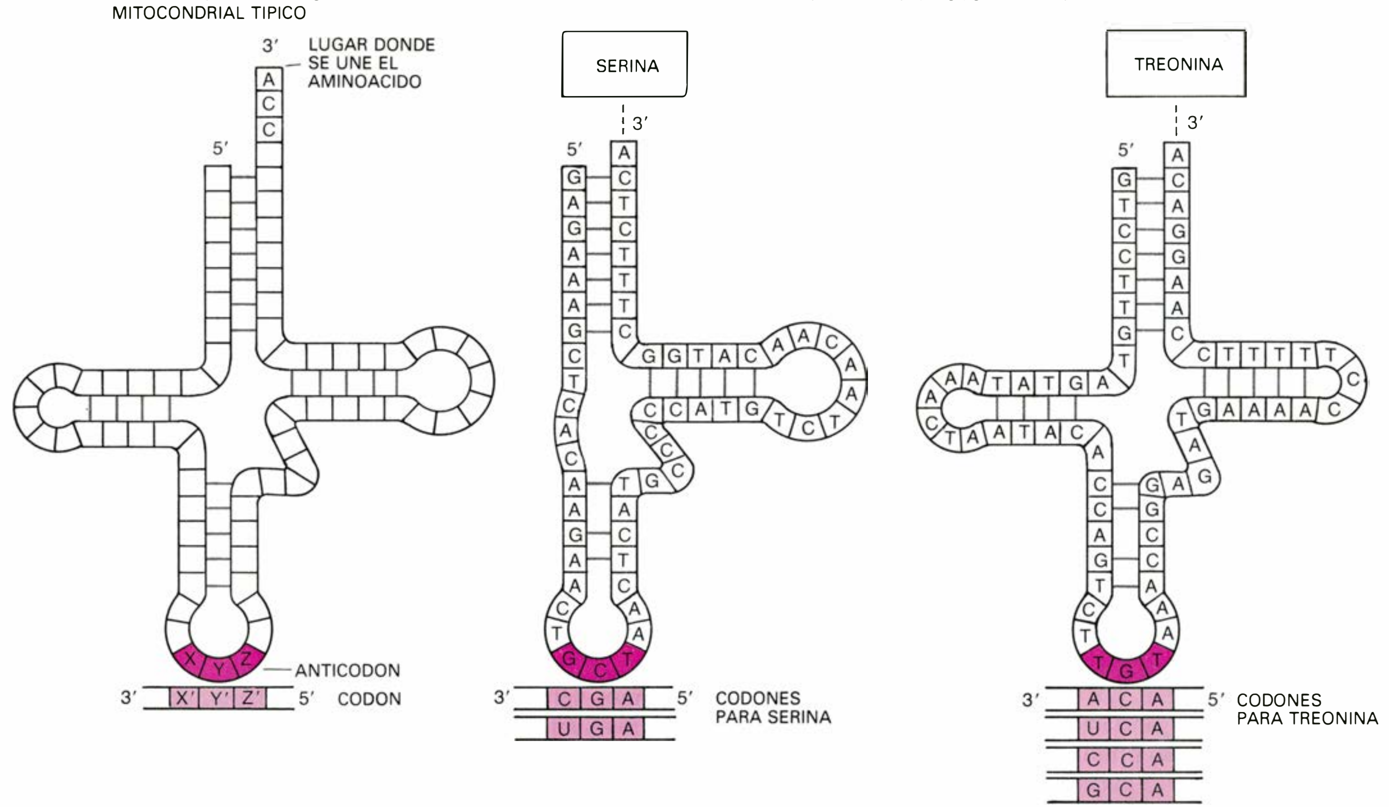
Walter Gilbert, de la Universidad de Harvard, sostiene que en los genomas víricos y nucleares los intrones pueden conferir una ventaja selectiva permitiendo una evolución más rápida. Aumentando la longitud de un gen se incrementa la probabilidad de que ocurran fenómenos de recombinación que permitan intercambiar segmentos análogos de dos alelos, o versiones alternativas de un mismo gen. Tal recombinación podría unir dos mutaciones ventajosas, cada una en un alelo diferente, creando así un gen doblemente ventajoso. También podría crear genes completamente nuevos mediante lo que Gilbert llama un "batido" o mezcla de exones: juntando, como exones de un solo gen, varias secuencias informativas que previamente determinasen proteínas diferentes. Es más, la mutación de una base situada en el límite entre un intrón y un exón podría cambiar el patrón de procesamiento y generar, en consecuencia, ARN mensajeros que determinasen proteínas muy diferentes. Los cambios en el procesamiento inducidos por tales mutaciones no tendrían por qué ser absolutos; el ADN afectado se convertiría en un gen "multiuso": podría determinar tanto la proteína original como la nueva, según adoptara un tipo de procesamiento u otro. El organismo podría ensayar una solución sin borrar de su memoria la antigua.

¿Cómo se originaron los intrones? La introducción de secuencias irrelevantes en un gen esencial suele ser letal. Aun cuando un mecanismo de procesamiento del ARN eliminase dichas secuencias, ¿por qué las fuerzas selectivas habrían de conservar un intrón recién introducido? La célula no se beneficia a corto plazo de la presencia de una secuencia intercalada de ese tipo, y ciertamente no puede predecir que vaya a ser ventajosa en el futuro. Consideraciones como estas llevaron a W. Ford Doolittle III, de la Universidad de Dalhousie en Nueva Escocia, a proponer que los intrones no se introducían en las células eucariotas, sino que formaban parte de su herencia evolutiva.

Según Doolittle, los procariotas actuales (las bacterias), así como los núcleos eucariotas, probablemente deriven de un primitivo organismo ancestral que fue capaz de evolucionar rápidamente porque disponía de una maquinaria para la replicación, transcrip-

PRIMERA LETRA	SEGUNDA LETRA				TERCERA LETRA
	U	C	A	G	
U	PHE	SER	TYR	CYS	U
	PHE	SER	TYR	CYS	C
	LEU	SER	*PARO*	*(PARO)* TRP	A
	LEU	SER	*PARO*	TRP	G
C	LEU	PRO	HIS	ARG	U
	LEU	PRO	HIS	ARG	C
	LEU	PRO	GLN	ARG	A
	LEU	PRO	GLN	ARG	G
A	ILE (MET)	THR	ASN	SER	U
	ILE	THR	ASN	SER	C
	(ILE) MET	THR	LYS	(ARG) *PARO*	A
	MET	THR	LYS	(ARG) *PARO*	G
G	VAL	ALA	ASP	GLY	U
	VAL	ALA	ASP	GLY	C
	VAL	ALA	GLU	GLY	A
	VAL	ALA	GLU	GLY	G

9. CODIGO GENETICO, en el que un triplete de nucleótidos (designados por sus iniciales *U, C, A* y *G*) es un codón que determina un aminoácido concreto o una señal terminadora. Se pensaba que el código era universal, pero varía en los sistemas genéticos mitocondriales. La tabla indica las desviaciones que ocurren en las mitocondrias humanas con respecto al código "universal"; las desviaciones se indican en letras de color, con la traducción "normal" entre paréntesis. En las mitocondrias humanas, los 22 ARN de transferencia "leen" dos (*recuadros grises*), o bien cuatro (*recuadros de color*) codones sinónimos. El codón correspondiente al ARN de transferencia para el aminoácido metionina no se conoce con seguridad.



10. LOS ARN DE TRANSFERENCIA presentan generalmente una estructura en forma de trébol (*izquierda*), que resulta del emparejamiento (*líneas grises*) entre bases complementarias. (En el ARN, *A* se empareja con *U* y *G* con *C*. En el ARN de transferencia mitocondrial que se indica aquí, la secuencia corresponde a la del ADN y no al ARN, por lo que *A* se empareja con *T*.) Muchos ARN de transferencia mitocondriales tienen una forma aberrante; al de la serina (*centro*) le falta todo un brazo. El anticodón, formado por tres nucleótidos, de cada uno de los ARN transferentes, lee un codón del ARN mensajero. En algunos casos, el ARN transferente puede leer más de un codón, ya que hay cierta libertad, o "tambaleo", en el emparejamiento entre la primera base de su anticodón y la tercera base del codón; la *X* del anticodón de la izquierda podría emparejarse con otra base, además de con *X'*. Muchos ARN de transferencia de mitocondrias humanas pueden leer hasta cuatro codones (*derecha*), de forma que con sólo 22 ARNt pueden leerse 61 codones.

ción y traducción un tanto chapucera e imprecisa. Para asegurar la correcta expresión de la información genética de un organismo como ese, dicha información tendría que haber sido redundante. La organización de los genes en fragmentos pudo haber sido esencial para asegurar que los productos de la transcripción de los exones se uniesen, al menos ocasionalmente, para dirigir la síntesis de proteínas funcionales. Cuando la maquinaria genética adquirió mayor precisión, sugiere Doolittle, tales medidas de seguridad se hicieron menos necesarias; la replicación y transcripción del ADN redundante y carente de información se tornó cada vez más gravosa, surgiendo la presión para eliminarlos. La eliminación supuso, como ventaja, un aumento de la eficacia, siendo un ejemplo de ello los perfeccionados genes de las modernas bacterias y los virus que las infectan. El precio pagado por esta mayor eficacia fue la pérdida de potencial para cambios evolutivos posteriores.

¿Qué ocurre con los intrones de los genes mitocondriales? Muchas de las ideas apuntadas para explicar los intrones en genes nucleares no son satisfactorias para los mitocondriales, no sólo porque no explican la existencia de las regiones potencialmente informativas, sino porque sus principales predicciones no se ven apoyadas por los hechos. La retención pasiva de intrones con regiones potencialmente informativas no es una buena explicación, ya que es apreciable el costo que le supone a la célula mantener y expresar todos los genes que hacen falta para el procesamiento de los intrones, y un sistema tan complejo debe ser muy vulnerable a las mutaciones. Que la retención sea para facilitar la recombinación es una explicación poco probable, ya que no hay indicios de que los nuevos genes mitocondriales surgieron recientemente por una mezcla de genes, o que las proteínas determinadas mitocondrialmente en los hongos estén evolucionando más rápidamente que las de mitocondrias de mamíferos, cuyos genes carecen de intrones. Por otra parte, si bien la mezcla de genes es, en principio, una forma excelente para propagar mutaciones que confieran características ventajosas (como resistencia a inhibidores naturales de la función mitocondrial) no hay razón para creer que tales mutaciones se fijen más eficazmente en los genes fragmentados que en los continuos. Finalmente, no hay indicios (si bien son escasos los datos incuestionables) de que alguno de los genes fragmentados de las mitocondrias de levaduras sean genes "multiuso".

Intrones y represión

Quizá, pues, los intrones con secuencias potencialmente informativas se hayan conservado en los ADN mitocondriales de hongos porque confieran algún tipo de ventaja selectiva. ¿Cuál podría ser esa ventaja? Según Piet Borst, de la Universidad de Amsterdam, tales intrones podrían desempeñar un papel en la regulación fina de la expresión de los genes mitocondriales. Las levaduras, cuando crecen por fermentación, en ausencia de oxígeno, pueden reprimir selectivamente la síntesis de ciertas proteínas mitocondriales, y parte de esa represión puede ejercerse a nivel del procesamiento del ARN. La propuesta de Borst tiene la ventaja de explicar por qué la represión es selectiva. Genes como los del citocromo *b* y el de la subunidad I de la oxidasa del citocromo *c*, con varios intrones cuya escisión está probablemente controlada por una madurasa, debe-

rían ser particularmente sensibles a la represión, y tal es lo que ocurre. El argumento de la represión, sin embargo, no explica la presencia de todos los intrones mitocondriales de levaduras. Conviene tener en cuenta, además, la posibilidad de que alguno de los intrones que carecen de secuencias potencialmente informativas no tenga función alguna. Si existe un mecanismo eficaz para escindirlos del ARN puede que exista poca presión selectiva para eliminarlos del ADN; de hecho, las mitocondrias pueden carecer de la maquinaria para hacerlo eficazmente. Los intrones podrían ser "ADN egoísta", que se autoperpetúa en interés propio.

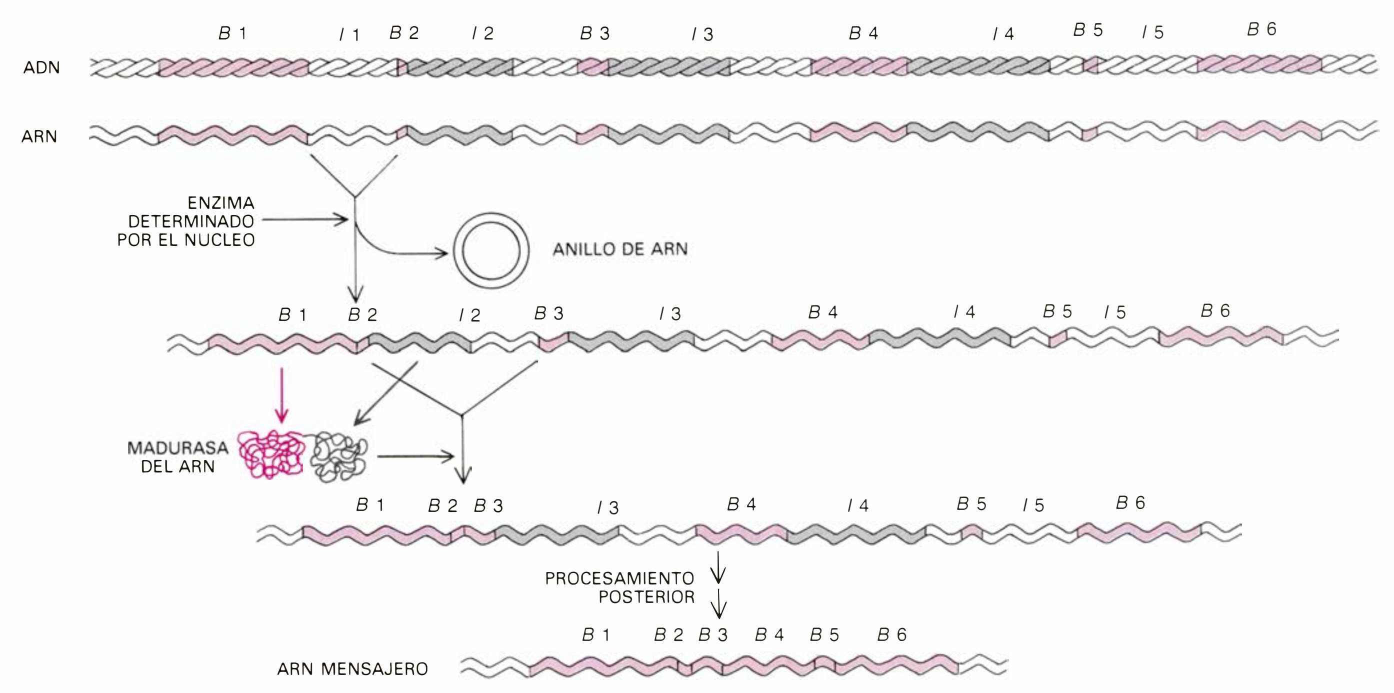
El hecho de que algunos intrones de mitocondrias de levaduras, *Aspergillus* y *Neurospora*, posean una disposición muy parecida dentro de los correspondientes genes fragmentados sugiere que el antecesor común de esos genes pudo haber tenido una organización similar. Esos intrones podrían, por ello, reflejar directamente la organización fragmentada de los genes del primitivo organismo endosimbiótico, a partir del cual, se piensa, han evolucionado las mitocondrias. No obstante, quizá no todos los intrones sean tan antiguos. Las comparaciones de las secuencias de intrones de levaduras realizadas, con fino detalle, por François Michel y Bernard Dujon, del Centro de Genética Molecular, y por Linda Bonen, de la Universidad de Amsterdam, revelan muchas similitudes en la estructura secundaria del ARN y en las secuencias con información para fabricar proteínas. Estas similitudes nos inducen a pensar que la mayoría de los intrones actuales, y sus secuencias potencialmente informativas, son miembros de una familia multigénica que se originó a partir de un único gen ancestral, que se expandió por duplicación y transposición.

Dicha duplicación con movilidad es característica del comportamiento de los transposones: segmentos de ADN que, en bacterias y en organismos superiores, replican e insertan sus copias aleatoriamente dentro de un genoma. Los intrones de levadura que muestran un gran parecido quizá deriven de transposones. Si éste es el caso, tendríamos ahí el origen de las madurasas. La mayoría de los transposones bacterianos determinan una transposasa: enzima que cataliza la inserción del transposón. La transposición implica cortes exactos en los dos extremos del transposón, como el procesamiento del ARN requiere cortes precisos en los extremos de un intrón. Por tanto, las madurasas podrían haber evolucionado a partir de las transposasas.

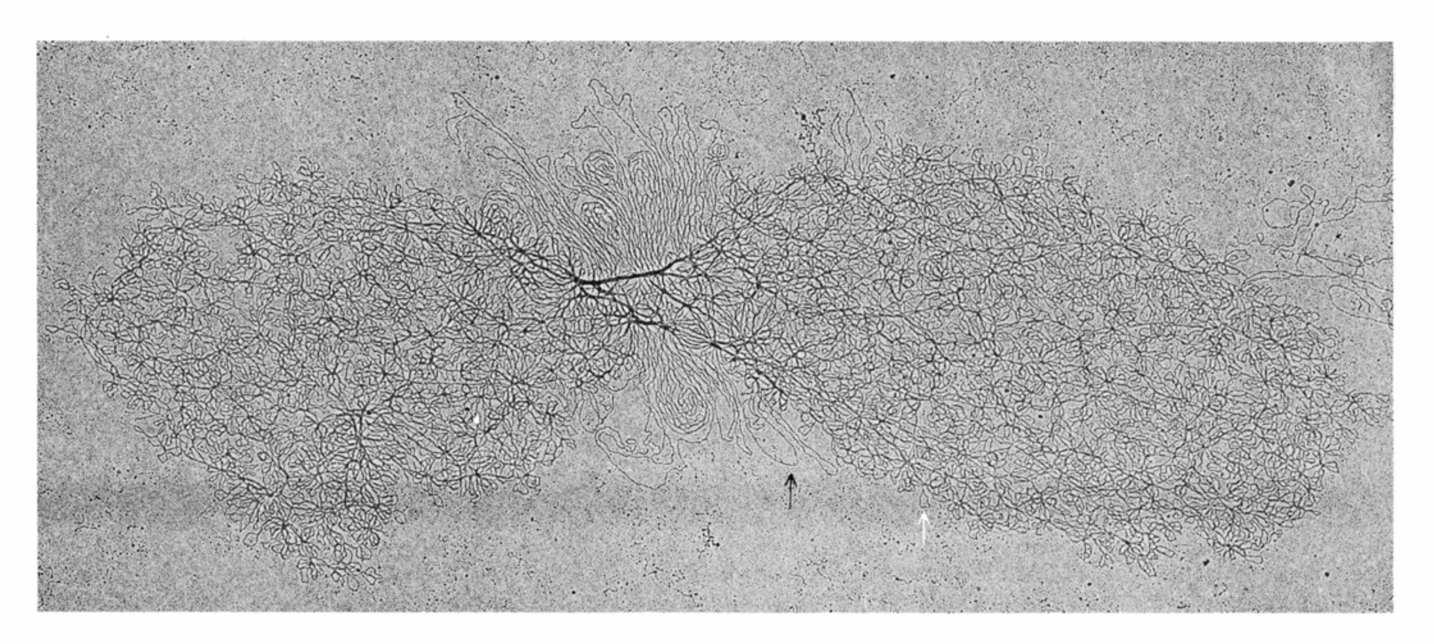
¿Por qué ADN mitocondrial?

¿Por qué es necesario mantener un sistema genético separado para uno solo de los muchos orgánulos celulares que existen? El sistema mitocondrial no puede explicarse como una forma de amplificación génica, necesaria para suministrar a la mitocondria grandes cantidades de ciertas proteínas. Hay otros medios de conseguir altos niveles de síntesis de proteínas, y si la amplificación fuese necesaria, bien podría tener lugar en el núcleo. Tampoco puede explicarse el ADN mitocondrial como fuente necesaria de proteínas cuya insolubilidad impidiera su transporte hasta su punto de ubicación, en la membrana interna. La subunidad 9 del complejo ATPasa, una de las proteínas más hidrófobas que fabrica, en levaduras, el ADN mitocondrial, está determinada por un gen nuclear en *Neurospora* y en mamíferos, en los que se sintetiza en los ribosomas del citoplasma y se introduce posteriormente en la mitocondria.

Si no se necesita un sistema genético mitocondrial distinto, ¿por qué está allí? Borst piensa que el ADN mitocondrial es, simplemente, un callejón evolutivo sin salida, una reliquia de la supuesta bacteria simbiótica. De acuerdo con la hipótesis de Borst, en las etapas iniciales de la evolución existían pocas barreras para la transferencia de información entre el sistema genético del endosimbionte y el del hospedador y, gradualmente, la mayoría de los genes del endosimbionte se fueron integrando en el genoma de la célula anfitriona. A medida que la célula eucariota ganaba en complejidad (su ADN se organi-



11. LA MADURASA DEL ARN, enzima determinado en parte por el segundo intrón del gen para el citocromo *b* en mitocondrias de levaduras, ayuda a catalizar la escisión del intrón que lleva la información para su síntesis. El gen tiene seis exones (de *B*1 a *B*6) y cinco intrones. (Los exones e intrones no se han dibujado a la misma escala.) Dicho gen se transcribe a ARN e inmediatamente es eliminado el primer intrón, gracias a un enzima procesador determinado por el núcleo. (Sorprendentemente, se acumulan copias del intrón en forma de anillos.) La escisión permite unir los exones *B*1 y *B*2 a una larga secuencia potencialmente informativa (*gris*) del intrón *I*2. Este ARN sirve de ARN mensajero para la síntesis de la madurasa, necesaria para eliminar *I*2; de esta forma, la madurasa controla su propia síntesis. Hay pruebas de que otras madurasas colaboran también en la escisión de los intrones *I*3 e *I*4.



12. ADN MITOCONDRIAL ATIPICO del protozoo *Trypanosoma brucei*, aumentado unas 16.000 veces en esta electromicrografía realizada por J. H. Hoeimakers, de la Universidad de Amsterdam. El ADN aparece aquí en el transcurso de su segregación: una vez replicado se divide para formar dos genomas durante la segmentación del orgánulo. El ADN es una red de maxicírculos (*flecha negra*) y minicírculos (*flecha blanca*) entrelazados. Los maxicírculos parecen equivaler al ADN mitocondrial de otros organismos; se desconoce en la actualidad la función que puedan desempeñar los minicírculos.

zó en cromosomas y quedó restringido al núcleo), y quizás debido a que los códigos genéticos mitocondriales y nucleares divergieron, decreció la posibilidad de transferencia; finalmente, se hizo imposible. Unos pocos genes quedaron aislados en las mitocondrias, y persisten porque son esenciales.

La hipótesis es de una sugestiva sencillez. Explica perfectamente que varios enzimas mitocondriales determinados por el núcleo tengan similitudes con los enzimas correspondientes en bacterias. Predice que algunos genes se encuentran en el ADN mitocondrial de un organismo y en el ADN nuclear de otro, como es el caso del gen para la subunidad 9 de la ATPasa. Permite situaciones intermedias, en las que estén presentes en la misma célula las versiones mitocondriales y nucleares del mismo gen. Esto puede ser verdad en *Neurospora*, donde Paul van den Boogaart y Étienne Agsteribbe, de la Universidad estatal de Groningen, han encontrado copias del gen para la subunidad 9 de la ATPasa tanto en ADN nuclear como mitocondrial.

La hipótesis de Borst predice, además, que algunos ADN mitocondriales pueden tener más genes de los que están presentes, como serie básica, en mitocondrias de levaduras y mamíferos. Las mitocondrias de plantas superiores tienen genomas de extraordinaria magnitud: desde los 250.000 nucleótidos del maíz hasta los 2,4 millones de la familia de la calabaza y el melón. Gran parte de las diferencias pueden deberse a la presencia de ADN no informativo, aunque algunos experimentos demuestran que las mitocondrias de plantas sintetizan más proteínas que las de hongos o mamíferos, y hay pruebas de que, en plantas, el fenómeno de la androesterilidad (incapacidad para producir polen fértil) obedezca a cambios en el ADN mitocondrial.

Evolución hacia la diversidad

Los principales productos de los genes mitocondriales parecen, en gran medida, haberse conservado a lo largo de la evolución; no ha ocurrido lo mismo con la organización y modo de expresarse de esos genes. Donde mejor se observa la diversidad de organización es en el ADN mitocondrial de *Trypanosoma*, un género de protozoos unicelulares que comprende parásitos del hombre y de animales domésticos. Su ADN es una enorme red de círculos enlazados, unos 10.000 quizá. Hay maxicírculos, que oscilan entre 20.000 y 36.000 nucleótidos, que se transcriben y poseen regiones homólogas a fragmentos del ADN de levadura; son, probablemente, el equivalente al ADN mitocondrial en otros organismos. La mayor parte del ADN, sin embargo, está formada por minicírculos de menos de 3000 nucleótidos. Estos no se transcriben, y han experimentado una alta tasa de evolución, característica del ADN no informativo. Se desconoce la función de la amplia red de minicírculos. Y tampoco se sabe por qué el ADN mitocondrial debe tener esa estructura tan poco corriente.

Espero que esta visita final a una de las avanzadillas de la biología convenza al lector de la diversidad de los sistemas genéticos mitocondriales. Variedad que resulta de una alta tasa de evolución en la secuencia de nucleótidos, quizá por encima de 10 veces la tasa de los genes nucleares. La aparente falta de limitación en la evolución de la secuencia puede guardar relación con la simplicidad del sistema. Tiene por función primaria sintetizar un número limitado de proteínas en cantidades muy constantes y es, además, capaz de alojar mutaciones que resultarían deletéreas en un sistema sujeto a la regulación diferencial de muchos genes. Cabe también la posibilidad de que, en cada organismo, diferentes presiones selectivas hayan empujado a los ADN mitocondriales a distintas y, a veces, muy extremas soluciones al problema de la expresión génica.

Queda para el futuro dilucidar cuáles sean estas presiones; al igual que averiguar, con el estudio de este sistema limitado, cuáles son los requerimientos mínimos de un sistema genético funcional. Un mayor conocimiento del genoma mitocondrial revelará algunas de las pautas evolutivas que han desembocado en los mecanismos de expresión génica de las bacterias actuales y de los eucariotas. A juzgar por lo que ya ha revelado el estudio de las mitocondrias, las enseñanzas que se obtengan sobre la organización y expresión del ADN mitocondrial se podrán aplicar mucho más allá de los límites de la membrana de este modesto orgánulo.

El futuro del universo

Predicción para el universo en expansión hasta el año 10^{100}: todos los protones se desintegrarán, las galaxias formarán agujeros negros y éstos se "evaporarán". Si el universo se colapsara, podría reaparecer e iniciar entonces un nuevo ciclo

Duane A. Dicus, John R. Letaw, Doris C. Teplitz y Vigdor L. Teplitz

El desarrollo de la física de interacciones a altas energías de las partículas elementales registrado a lo largo de los últimos años ha hecho avanzar con firmeza la cosmología. El esfuerzo puesto en hallar una sola fuerza que explicara las cuatro fundamentales de la naturaleza como aspectos distintos suyos ha cosechado éxitos parciales con las "teorías de gran unificación" (TGU) de las partículas elementales. Armados con esas teorías, podemos dar una descripción aproximada de los procesos físicos más importantes desde temperaturas extremadamente bajas hasta temperturas del orden de 10^{32} grados Kelvin. Lo que nos abre el camino para abordar las propiedades de la materia con densidades de interés cosmológico que van desde 10^{-300} gramos hasta más de 10^{100} gramos por centímetro cúbico. Las desiguales condiciones que se presentan en los extremos de esta escala tienen importancia sólo en los primerísimos instantes y en los postreros de la evolución del universo.

Desde hace unos años, no muchos, los cosmólogos empezaron a ver que la física teórica podía explicar, dentro de un marco coherente, la historia del universo desde sus primeros 10^{-35} segundos de existencia. Más recientemente, algunos físicos y cosmólogos, entre los que nos incluimos los cuatro firmantes, han extrapolado los episodios cósmicos hacia un futuro que se aleja hasta unas 10^{100} veces la edad actual del universo. Varios son los motivos para acometer el estudio, aparte de la fascinación intrínseca que conlleva. Desde el punto de vista de la física teórica, dicha extrapolación permitiría un experimento idealizado: examinando en un instante dado cómo interaccionan los distintos efectos predichos por una teoría, podría verificarse la coherencia global y plausibilidad de dicha teoría. Ningún laboratorio terrestre puede alcanzar, siquiera fuera aproximadamente, las temperaturas y densidades para las que se hacen predicciones en las teorías de gran unificación, por lo que éstas se verifican en lo que algunos físicos llaman el laboratorio fundamental, o sea, el universo considerado como un todo. Desde el punto de vista de la cosmología, la importancia de nuestras extrapolaciones se basa en que las teorías de gran unificación tienen consecuencias que pueden comprobarse en laboratorios terrestres y podrían así confirmarse dichas predicciones referentes a condiciones extremas. Nos es permitido, pues, describir una serie de acontecimientos de un futuro muy lejano con un detalle mucho más fino de lo que hasta hace poco era posible.

El marco de nuestros cálculos de este futuro lejano es el modelo de la gran explosión, según el cual el universo empezó con la explosión de una entidad primordial extremadamente compacta, 10.000 o 20.000 millones de años atrás. La expresión "gran explosión" y la terminología asociada es, por una parte, adecuada ya que, desde un origen común, se lanzó materia y energía hacia el espacio exterior. Pero también dicha expresión puede inducir a error, ya que no es correcto suponer que la explosión pueda ser vista desde fuera. No hay ningún punto de vista exterior, ya que lo que está explotando es todo el universo. El mismo espacio está en expansión, en el sentido de que la separación entre dos galaxias cualesquiera crece con una rapidez que depende de esta separación. Desde nuestra galaxia parece que todas las demás se alejan con una rapidez que aumenta en 17 kilómetros por segundo cada millón de años-luz de separación. La teoría de la relatividad general de Einstein describe, desde un punto de vista matemático, la gran explosión.

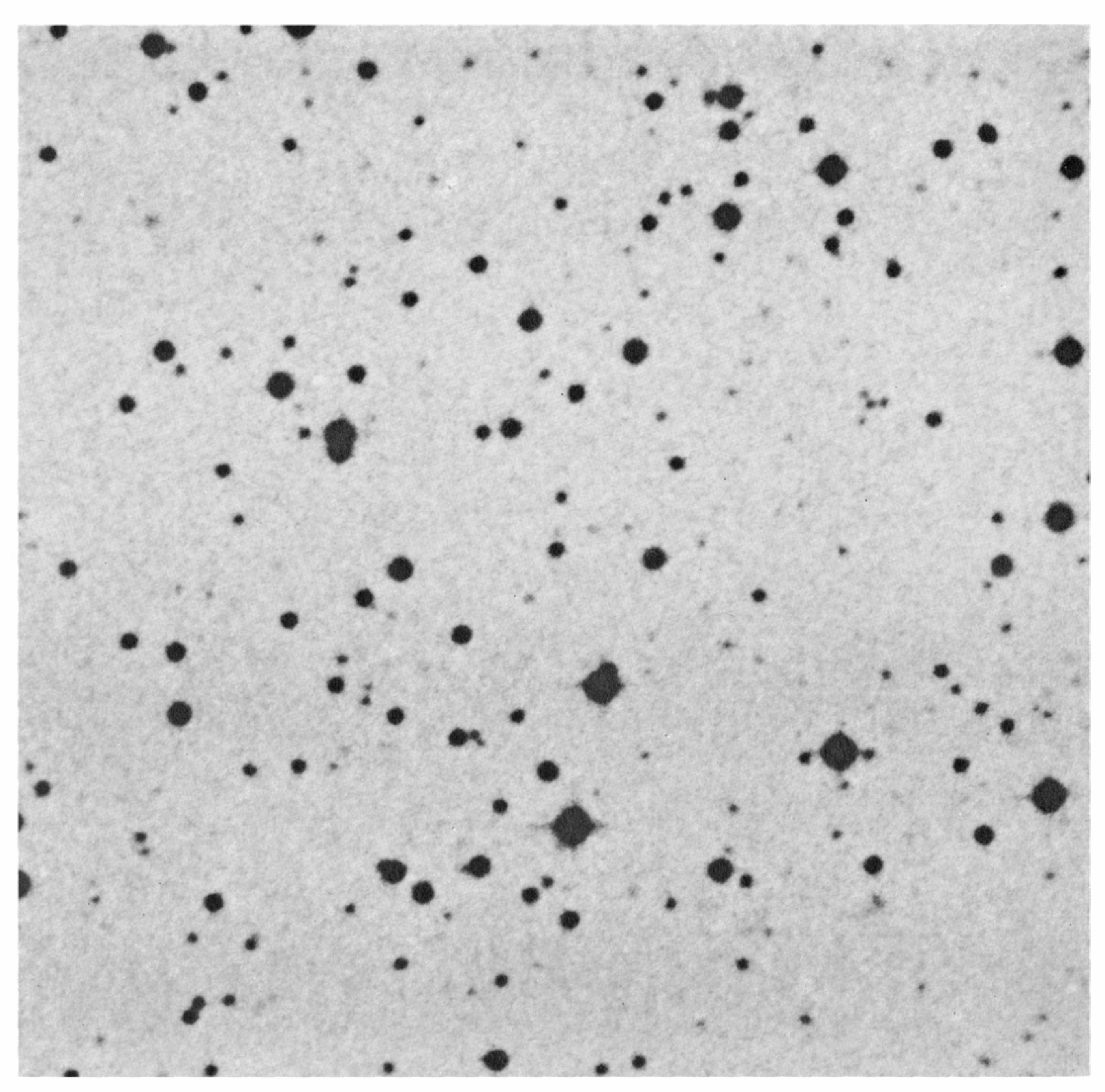
El distanciamiento de las galaxias lejanas se deduce del corrimiento hacia el rojo que observamos en el espectro electromagnético de sus radiaciones. La expansión y enfriamiento del universo ha afectado mucho más drásticamente al espectro de la otra clase de radiación, la llamada radiación cósmica de fondo en microondas, descubierta en 1964 por Arno A. Penzias y Robert W. Wilson, de los Laboratorios Bell. Al parecer, la radiación llega con la misma intensidad desde cualquier punto del universo; su intensidad varía con la dirección en menos de una parte sobre diez mil. La causa que produjo la radiación de fondo fue la difusión y aniquilación de partículas cargadas eléctricamente, en las primeras etapas de la gran explosión. Por tanto, en aquel momento la radiación era muy caliente. Desde entonces, y debido a la expansión del universo, se ha ido enfriando hasta una temperatura de unos tres grados Kelvin.

Las reacciones nucleares que tuvieron lugar con las densidades y temperaturas correspondientes a los tres primeros minutos subsiguientes a la gran explosión son las responsables de la síntesis del helio y también, aunque en mucha menor parte, de otros elementos ligeros. Pero ni el carbono ni los elementos más pesados pudieron formarse, ya que el universo se enfrió muy deprisa, y quedó una importante cantidad de hidrógeno que sirve de combustible nuclear para las estrellas. El exceso observado de materia sobre antimateria pudo ser causado, muy tempranamente, en las reacciones nucleares ocurridas en los 10^{-38} segundos después del inicio del universo; en cuyo caso, la mayoría de las teorías de gran unificación pronostican la desintegración final de toda la materia nuclear. Nuestros conocimientos sobre ese universo incipiente no nos faculta todavía para dar respuesta a la pregunta más apremiante que inquieta a la cosmología: ¿Continuará indefinidamente la expansión del universo o detendrán su curso las fuerzas gravitatorias hasta el punto de in-

vertir el sentido arrastrando de nuevo todo el espacio y el tiempo hacia la bola de fuego original?

Tras múltiples y reiteradas observaciones y reflexiones teóricas empeñadas estos años en resolver la cuestión, ésta sigue todavía abierta. En principio, la forma más inmediata de abordarla es midiendo la cantidad de materia que hay en un volumen grande y representativo del espacio y, a partir de ahí, deducir la densidad que tiene el universo a gran escala. Si la densidad es igual o mayor que una densidad crítica, la que darían unos tres protones por cada 1000 litros de espacio, la gravedad podría finalmente contener la expansión y el universo se colapsaría. Un universo que se comportara así sería un universo cerrado. Si la densidad es menor que el valor crítico, el universo permanecería abierto; la gravedad frenaría la expansión, pero nunca la detendría. Conviene señalar que ni la combustión estelar ni la desintegración de partículas elementales, u otros procesos físicos de este tipo, pueden transformar un universo cerrado en otro universo abierto.

Las estimaciones referentes al número de estrellas de una galaxia normal y las del número de galaxias por unidad de volumen dan un valor, para la densidad de la materia, que es sólo el 5 por ciento de la densidad crítica. Según esos resultados, podría concluirse que el universo es abierto y continuaría expansionándose indefinidamente. El cálculo de la masa fundado en el recuento de estrellas sólo incluye la materia luminosa; pero si la materia no luminosa

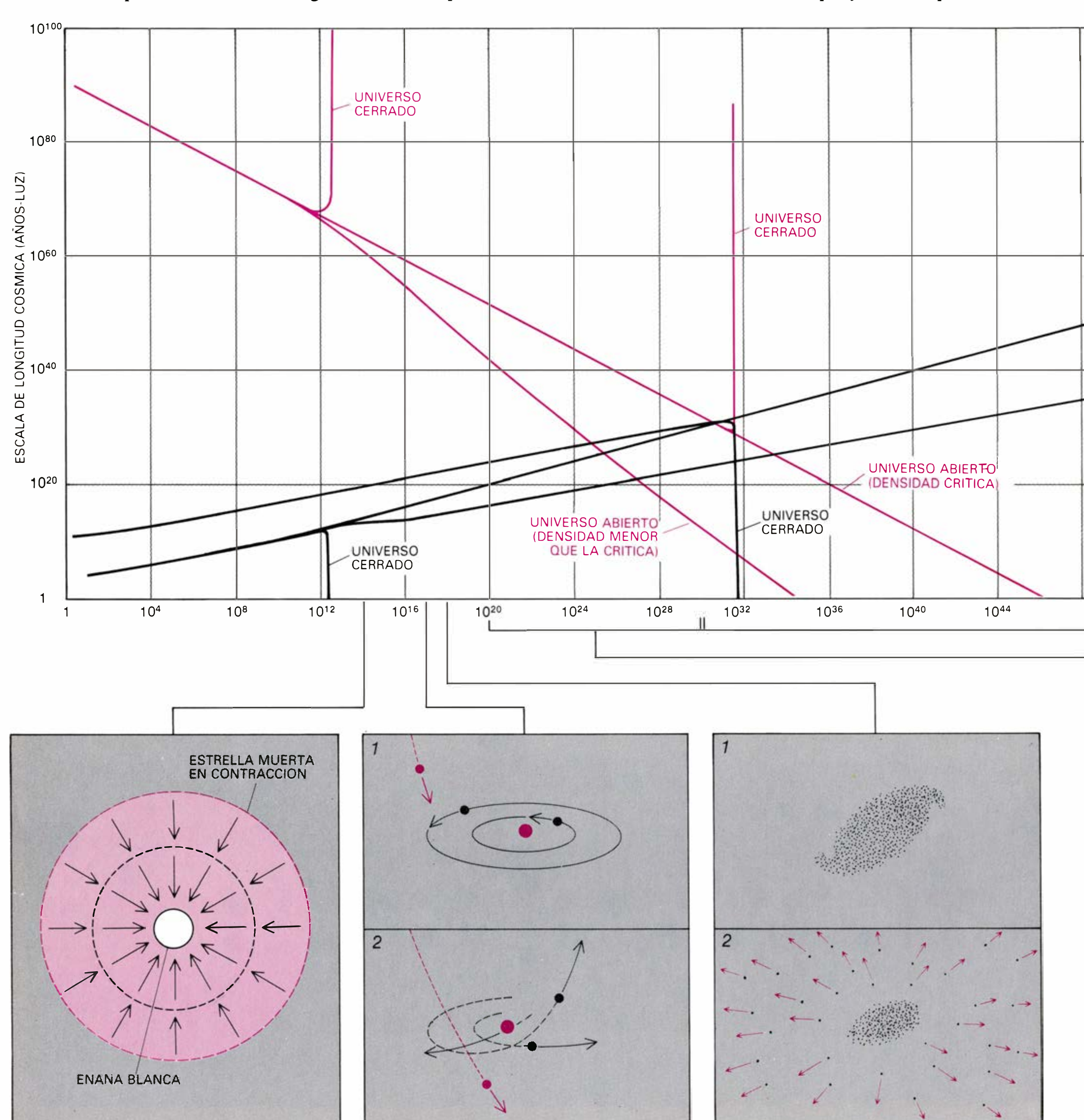


1. OBJETO MAS DISTANTE CONOCIDO: el quásar designado por PKS 2000-330. Su corrimiento hacia el rojo es de 3,78, lo que significa que está alejándose de nosotros con una velocidad igual al 92 por ciento de la velocidad de la luz. Esta enorme velocidad de recesión viene provocada, probablemente, por la gran explosión; si ello es cierto, PKS 2000-330 se halla a más de 16.000 millones de años-luz de nuestra galaxia. La fotografía es un negativo; el quásar es el objeto oscuro en forma de estrella que ocupa el centro geométrico. Estos objetos tan lejanos permiten deducir qué condiciones reinaron en el remoto pasado y así extrapolar ciertos sucesos hacia el futuro. La luz del quásar que ahora nos llega se emitió poco después de la gran explosión.

constituyera una porción importante de la masa del universo, semejante estimación resultaría inadecuada.

Tenemos datos que ponen de manifiesto la existencia de materia no luminosa en las galaxias. La velocidad con que una estrella gira en torno al centro de su galaxia debería crecer al aumentar la distancia a dicho centro, para las estrellas interiores a la galaxia; y decrecer respecto de dicha distancia, si las estrellas son periféricas. Sin embargo, para las estrellas exteriores no se observa este esperado decrecimiento de la velocidad de giro; la distribución de velocidades que se aprecia nos induce a concluir que, en la periferia de la galaxia, hay un halo de materia no luminosa cuya masa total es una o dos veces mayor que la de la materia luminosa de la galaxia: la materia no luminosa contenida en el interior galáctico contribuiría con un 5 o 10 por ciento a la masa necesaria para cerrar el universo.

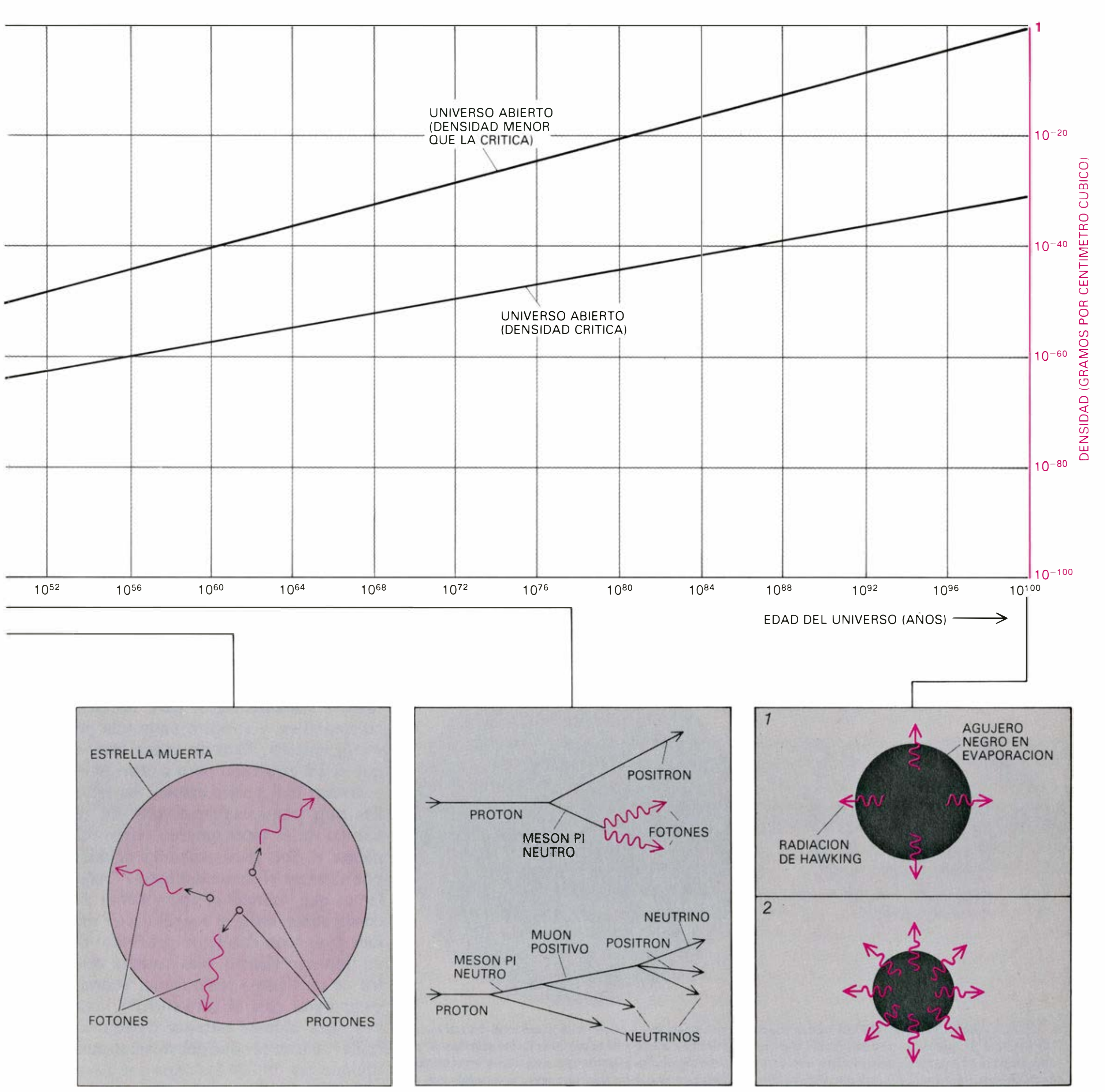
De un modo semejante, de la observación de los movimientos de las galaxias, dentro de un cúmulo, parece deducirse la existencia de materia no luminosa en una cantidad que daría el 50 por ciento de la densidad crítica del universo. La materia no luminosa explicaría por qué las galaxias se agrupan en cúmulos, en vez de moverse en direcciones arbitrarias. La materia oscura puede hallarse en forma de rocas frías, de estrellas que ya se han quemado, de



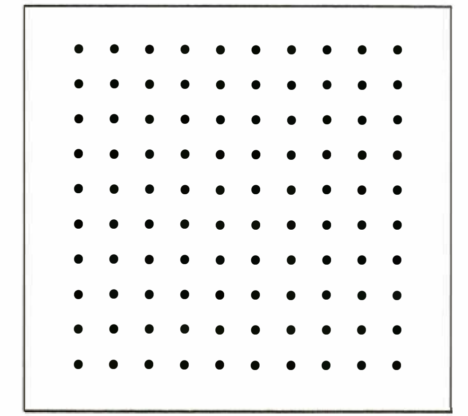
2. SE ESPERA QUE SE DEN SEIS GRANDES ETAPAS en el futuro del universo, de acuerdo con las teorías físicas y a condición de que la expansión del mismo iniciada en la gran explosión dure lo suficiente para que puedan desarrollarse. Aquí se dibuja la gráfica de la dependencia de la escala de longitud (que es más o menos equivalente al radio del universo) y la densidad respecto del tiempo para cuatro modelos de universo: uno corresponde a un universo abierto, dos a universos cerrados y uno en la frontera, definida matemáticamente, entre los modelos cerrados y los completamente abiertos. En un universo abierto, la densidad de materia y energía es demasiado baja para que la gravedad domine la expansión y, por tanto, el universo se expansionaría indefinidamente. Si el universo es cerrado, la gravedad acabará por detener la expansión y arrastrará al universo de nuevo hacia un punto de densidad infinita. No se sabe si estamos en un universo abierto o en otro cerrado. Las etapas futuras aparecen esquematizadas. Las estrellas agotarán su com-

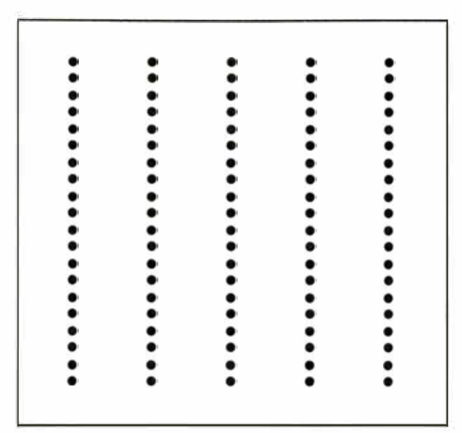
nubes de gas o incluso de agujeros negros.

Si existiera actualmente la materia adicional para alcanzar el 50 por ciento que falta para llegar a la densidad crítica, podría encerrarse en forma de átomos de hidrógeno ionizados o en neutrinos. Los telescopios de rayos X en órbita alrededor de la Tierra deberían detectar el hidrógeno ionizado. En cuanto a los neutrinos, habría los mismos que fotones, o cuantos de energía electromagnética, hay asociados a la radiación cósmica de fondo; lo que representa varios centenares de neutrinos por centímetro cúbico. Su contribución a la masa del universo depende de la masa del neutrino. Se suponía que la masa del neutrino era nula, pero los experimentos actuales sólo dan una cota superior para la masa del neutrino y no eliminan la posibilidad de atribuirle cierta masa a esta partícula, suficiente para cerrar el universo. Además, se han aducido ya razones teóricas que abogan por la masa del neutrino, y son varios los grupos de experimentadores empeñados en detectarla y medirla. A lo largo de los próximos años estos experimentos podrían responder definitivamente, en un sentido afirmativo, a la pregunta del cierre. Por otro lado, un resultado negativo dejaría la cuestión pendiente de una mayor precisión en las medidas de las diferencias y veloci-



bustible y colapsarán en virtud de su propio peso dentro de 10^{14} años; las estrellas perderán sus planetas por choques con otras estrellas dentro de 10^{17} años. Las galaxias cederán gran parte de su masa dentro de unos 10^{18} años debido a encuentros todavía más próximos con otras estrellas y a causa de la difusión de átomos de hidrógeno y de partículas de polvo; aquellos conjuntos de materia que en los choques adquieran gran velocidad vencerán a las fuerzas gravitatorias internas de la galaxia, y la materia restante colapsará dando un agujero negro de gran masa. (Un agujero negro es una región del espacio de donde nada puede escaparse, salvo por mecanismos que describe la teoría cuántica.) Dentro de las estrellas frías, las que han escapado de las galaxias, se genera calor por desintegración de los protones. Este calentamiento será significativo hacia el año 10^{20}; para el año 10^{30}, un 40 por ciento del universo habrá sufrido esta desintegración. Alcanzados los 10^{100} años, los agujeros negros de gran masa la perderán a través de la "evaporación" cuántica.





3. DISTRIBUCION DE MATERIA Y ENERGIA DEL UNIVERSO; parece, con un alto grado de precisión, homogénea e isótropa. En un universo homogéneo, en volúmenes comparables del espacio hallamos el mismo número de galaxias. En un universo isótropo, además de la homogeneidad, la densidad de materia y energía es la misma en todas direcciones. Las dos propiedades no son equivalentes: un universo homogéneo no tiene por qué ser también isótropo. Si la materia está distribuida tal como se indica en el diagrama de la izquierda, el universo será, a la vez, homogéneo e isótropo. La distribución dibujada a la derecha es homogénea, aunque no isótropa: la distancia entre galaxias que están en la dirección vertical es menor que la distancia entre las que están en dirección horizontal. Se puede producir anisotropía en un universo homogéneo en expansión en el caso de que la intensidad de expansión varía con la dirección.

dades de recesión de las galaxias lejanas.

Aunque nuestra falta de certeza sobre la densidad real del universo haga que el valor posible oscile desde un universo cerrado hasta otro abierto, dicha oscilación no es mayor que dos órdenes de magnitud. El hecho de que el universo, en el momento presente de su historia, no haya iniciado la fase de concentración indica que su densidad no supera, en más de diez veces, el valor crítico. La estimación de la densidad del universo contando sólo la materia luminosa manifiesta que la densidad actual debe ser, como mínimo, la veinteava parte de la densidad crítica. Por ahora, nadie sabe explicar esto, que, en opinión de los cosmólogos, constituye una sorprendente coincidencia.

Ya que los resultados experimentales o de las observaciones no aclaran todavía la cuestión del cierre, debemos considerar las dos posibilidades –abierto y cerrado– para cualquier lucubración relativa al futuro lejano. Empecemos por suponer que no se alcanza la densidad crítica y el universo es abierto. ¿Qué sucederá con mayor verosimilitud? La pregunta debe contestarse tanto para la estructura a gran escala (esto es, indicar cómo evolucionan las propiedades geométricas del universo) cuanto para la composición local (esto es, para conjuntos de materia cuyo tamaño va desde el protónico hasta el galáctico).

La composición local del universo pasará por seis transiciones principales. En primer lugar, transcurridos 10^{14} años desde la gran explosión, todas las estrellas se habrán quedado sin combustible. El principal combustible nuclear es el hidrógeno, que se convierte en helio en el núcleo de la estrella durante la mayor parte de su vida. Cuando se ha gastado la mayor parte del hidrógeno combustible, la estrella se dilata rápidamente hasta alcanzar un tamaño varias veces superior al inicial y se transforma en una gigante roja; en esta etapa de la vida de la estrella, normalmente el helio se transforma en carbono y elementos más pesados. Esas reacciones nucleares son transiciones en un solo sentido: el hidrógeno se convierte en helio, el helio en carbono y el carbono en elementos más pesados, en una serie que suele acabar en el hierro. La energía total por unidad de masa del hierro es la más baja de todos los núcleos, de manera que, cuando se llega al "límite del hierro", toda la energía del universo que está almacenada como combustible nuclear se habrá consumido.

La intensidad con que una estrella consume su combustible nuclear depende de la masa de la misma: a mayor masa, más rápida se quema una estrella y, por tanto, más corta es su vida. Por ejemplo, el Sol habrá utilizado casi todo su hidrógeno dentro de unos 10.000 millones de años, después de lo cual variará rápidamente de tamaño, quemará algunos de los elementos más pesados con una asombrosa intensidad y en un tiempo bastante corto, para terminar colapsándose y constituyendo una pequeña estrella, llamada enana blanca, que se irá enfriando poco a poco. Para atravesar todas estas etapas, las estrellas más pequeñas precisarán de un tiempo varias veces superior al que empleara el Sol, pero acabarán también por alcanzar el límite del hierro. Señalemos que, aunque el agotamiento del combustible nuclear constituya el primer gran acontecimiento del futuro de un universo abierto, ello ocurrirá dentro de un tiempo remotísimo, enorme comparado con el transcurrido hasta hoy. Las últimas estrellas dejarán de brillar cuando la edad del universo multiplique por 10.000 los años que llevamos desde la gran explosión.

En una segunda etapa, las estrellas perderán sus planetas. Si una estrella se acerca a otra que tiene un pla-

4. LA MATERIA NO LUMINOSA puede constituir una parte importante de la masa total del universo. Hay pruebas de que esta masa integraría al menos la mitad de la masa de la mayoría de las galaxias. Si en éstas hubiera sólo materia que pudiera observarse directamente, las estrellas que estuvieran en la periferia de la galaxia girarían alrededor del centro galáctico con una velocidad que iría decreciendo con la distancia al centro (*ilustración de la izquierda*). Sin embargo, la velocidad que se observa de las estrellas periféricas no remite con la distancia. Semejante discrepancia encuentra explicación admitiendo la existencia de materia no luminosa distribuida por toda la galaxia y en un halo que la rodea. Por tanto, las estrellas que están en la periferia se hallarían en medio de la materia galáctica y su distribución teórica de velocidades respecto del centro de la galaxia se acercaría más a la distribución observada (*derecha*). Las longitudes de las flechas de uno y otro diagrama son proporcionales a las velocidades de las estrellas.

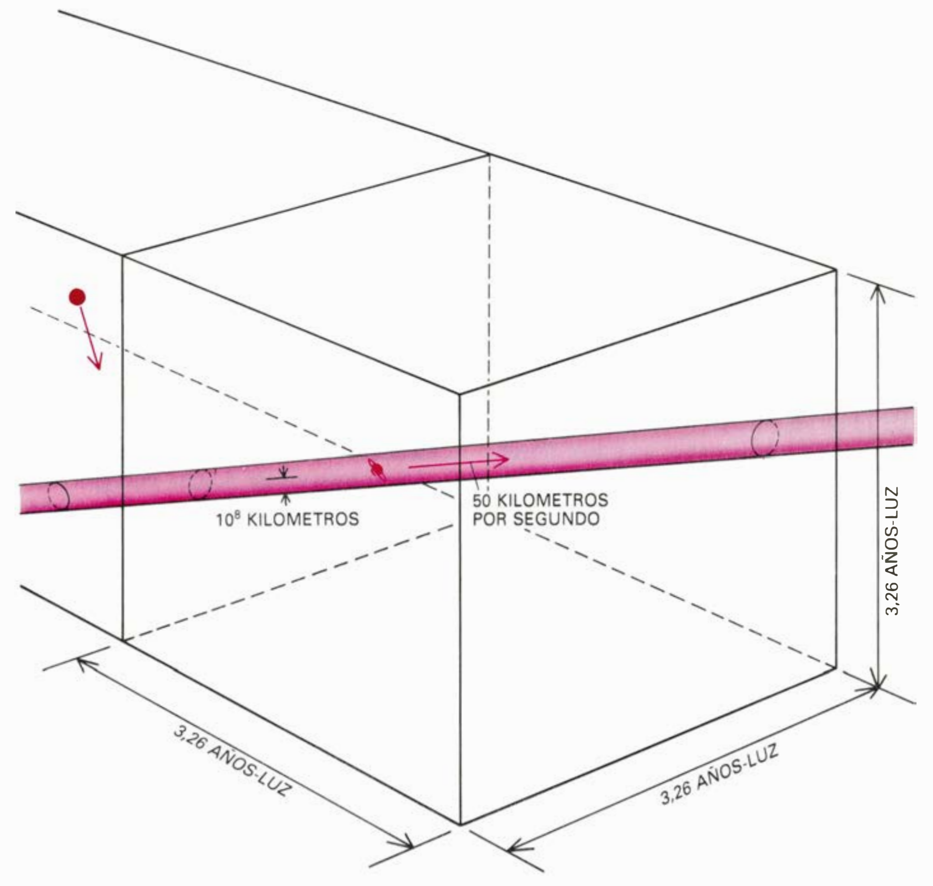
neta, entrando dentro del radio de la órbita planetaria, dicha órbita sufrirá un brusco cambio debido al campo gravitatorio de la estrella que pasa y el planeta puede verse despedido hacia el espacio. El tiempo promedio que transcurrirá antes de que ocurra un encuentro de este tipo dependerá de la densidad de estrellas de una región determinada, del área de las órbitas planetarias y de la velocidad relativa que hay entre las estrellas. Podemos expresar la densidad de las estrellas dando el volumen en el cual probabilísticamente debe haber una por lo menos. La estrella que se mueve por el espacio con un planeta traza un volumen cilíndrico cuyo tamaño depende del área de la órbita planetaria y de la velocidad de la estrella. El tiempo medio entre dos encuentros estelares es igual al tiempo que se necesita para que el volumen del cilindro sea igual al volumen donde puede hallarse una estrella [*véase la figura 5*].

En una galaxia, la densidad de estrellas es de una por cada cubo de 35 años-luz de lado. Freeman J. Dyson, del Instituto de Estudios Avanzados de Princeton, sugiere que un valor razonable para una órbita planetaria es de unos 100 millones de kilómetros, y añade que una estrella normal arrastra sus planetas por el espacio a unos 50 kilómetros por segundo. El volumen del cilindro que traza el sistema planetaria alcanzará, pues, los 35 años-luz cúbicos al cabo de unos 10^{15} años; a lo largo de ese intervalo es probable un encuentro capaz de provocar perturbaciones. Puede afirmarse sin miedo a errar que, al cabo de unos cien encuentros de esta clase, todos los planetas de estrellas habrán salido de su órbita; o sea: dentro de unas cien veces 10^{15} años, o 10^{17} años, las estrellas habrán perdido sus planetas.

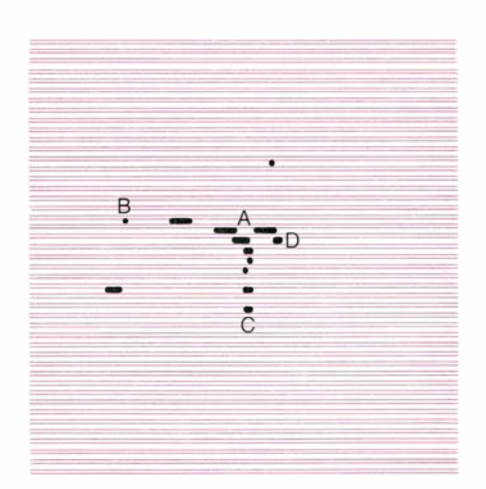
La tercera transición prevista es el resultado de encuentros estelares todavía más cercanos; a una escala galáctica sus efectos son evidentes. Cuando dos estrellas se aproximan, la interacción gravitatoria transfiere energía cinética de una a la otra. Si la aproximación aumenta, una de las estrellas puede ganar tanta energía que alcance la velocidad necesaria para escapar de la galaxia. Como en la interacción se conserva la energía, la otra estrella debe perder energía cinética y, por tanto, quedar más fuertemente ligada al núcleo de la galaxia.

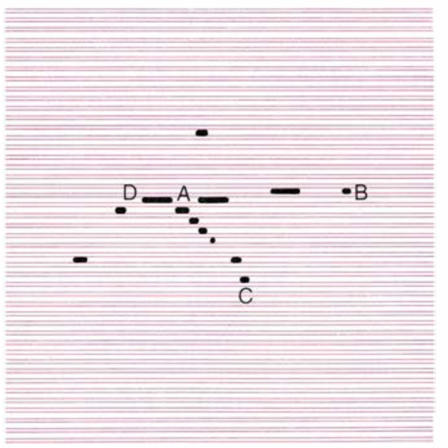
Podemos llamar a este proceso evaporación galáctica, ya que las interacciones entre las estrellas reproducen, a una escala grandiosa, las interacciones entre moléculas que se evaporan en la superficie de un líquido. En un intercambio de energía parecido puede estribar también la razón del escape, de la galaxia, de una fracción importante de gas interestelar. Evaporado el noventa por ciento de la masa de la galaxia, el campo gravitatorio arrastrará las estrellas y el polvo restante hacia un núcleo cada vez más denso. Es muy probable que las galaxias encierren hoy en su centro un agujero negro de gran masa, una región del espacio de donde sea imposible escapar (excepto a través de mecanismos descritos por la teoría cuántica). Si no existiera ahora dicho agujero negro, el núcleo galáctico desarrollaría una densidad tal que la gravitación acabaría por vencer la resistencia ofrecida por la presión del gas y el núcleo finalmente colapsaría formando un agujero negro de gran masa. Un cálculo parecido al que hemos indicado para la pérdida de los planetas nos lleva a esta otra conclusión: la evaporación de las estrellas y el colapso de las galaxias terminaría dentro de 10^{18} años.

Las transiciones cuarta y quinta que prevemos para el universo abierto se refieren a los últimos efectos cosmológicos predichos por la mayoría de los modelos de gran unificación de la física de partículas, efectos cuya importancia comenzará a revelarse cuando el universo centuplique la edad que tendrá al colapsarse las galaxias. Las teorías de gran unificación se proponen presentar una amplia descripción de las fuerzas fuertes, débiles y electromagnéticas entre partículas elementales. En el rango relativamente bajo de energías que se alcanzan en los laboratorios terrestres, esas tres fuerzas se muestran del todo diferentes; lo que no empece que dichas teorías las describan como manifestación de una sola interacción cuya unidad se percibe a las energías que corresponden a una temperatura del orden de 10^{27} grados Kelvin. A tamaña temperatura, partículas del tipo de los "quarks", que son sensibles a la fuerza fuerte, pueden transformarse en partículas tales como el electrón y el posi-



5. NUMERO DE ENCUENTROS entre una estrella y un sistema planetario. Puede deducirse del número de estrellas que hay por unidad de volumen, del tamaño del sistema planetario y de su velocidad respecto de las demás estrellas. Los encuentros alteran la órbita planetaria cuando la estrella corta el volumen que traza la órbita del planeta. En media, un choque así tendrá lugar cuando transcurra un tiempo igual al que precisa la órbita planetaria para trazar un volumen igual al volumen de un cubo dentro del cual se espera encontrar una sola estrella. Hay una estrella en la galaxia por cada 35 años-luz cúbicos. Cada sistema planetario se mueve a unos 50 kilómetros por segundo; las órbitas de los planetas exteriores trazan cilindros cuyo radio es de unos 100 millones de kilómetros. Bajo estas hipótesis y mediante un cálculo fácil se deduce que cada 10^{15} años se produce un encuentro. Después de unos 100 episodios, para los que serían necesarios 10^{17} años, las órbitas estarán tan seriamente perturbadas que los planetas se habrán separado de su estrella madre y dispersado por el espacio interestelar. Se dibuja muy exagerada la relación entre el tamaño de las estrellas con sistema planetario y el cubo que contiene una estrella.





6. DESINTEGRACION DEL PROTON, un suceso posible que tendría importantes consecuencias para el futuro del universo. Dicha desintegración, si no es demasiado rara, debería observarse experimentalmente. La mayoría de las teorías actuales de "gran unificación" de las interacciones entre partículas elementales predicen que la vida media del protón está entre 10^{30} y 10^{32} años. Para detectar esta desintegración, se han desplegado 13 experimentos, montados ya o en fase de proyecto. Se están estudiando sucesos, como candidatos para esta desintegración, procedentes de dos experimentos. Algunos de estos sucesos se han detectado por investigadores que trabajan en la mina de oro de Kolar, cerca de Bangalore, en la India meridional; de otro suceso aportan información investigadores del Centro Europeo de Investigaciones Nucleares (CERN), que trabajan con un detector instalado en el túnel del Mont Blanc, entre Francia e Italia. Aquí se presentan dos esquemas del suceso del Mont Blanc, en ángulo recto el uno respecto del otro. Los planos paralelos son placas paralelas de hierro, controladas para las desintegraciones; los puntos negros representan los productos de la desintegración. Si se verificara la desintegración del protón, se establecería un hito en la evolución hacia el año 10^{30}. Por aquel entonces se habría desintegrado la materia de gran longevidad menos electrones, positrones, fotones, neutrinos y agujeros negros.

trón, que a energías más bajas sólo son sensibles a las fuerzas débiles y electromagnéticas. A temperaturas más bajas, estos cambios de identidad son muy poco probables, aunque ocurran ocasionalmente. El protón, del que se supone que constituye un sistema ligado de tres quarks, podría desintegrarse si los quarks integrantes experimentasen dicha transformación.

Según la mayoría de las teorías de gran unificación, todos los protones se desintegran dentro de un período de 10^{30} a 10^{32} años; debería, pues, observarse la desintegración de al menos un protón por año en un conjunto que tuviera 10^{32} protones, como sería una masa de agua que pesara unas 160 toneladas. Hay en la actualidad 13 experimentos, en marcha o en estudio, que pretenden controlar grandes masas de agua, hierro u otros materiales para poner en evidencia esta desintegración del protón. Hay indicios de haberse descubierto una desintegración en una masa de 150 toneladas, sometida a control a 2300 metros bajo tierra en las minas de oro de Kolar, cerca de Bangalore, en la India meridional. Los detectores instalados en el túnel del Mont Blanc, entre Francia e Italia, han registrado una señal indicadora de una posible desintegración. Es demasiado pronto para juzgar si habrá más señales de éstas, lo que daría pie para pensar seriamente en realizar un escrutinio.

Si el protón se desintegrara, produciría un efecto importante en las estrellas que se hubieran salvado de la captura por parte de un agujero negro galáctico. Nos referimos a las estrellas que se escapan por evaporación; la desintegración de sus protones y neutrones las conservarán considerablemente más calientes que el medio interestelar que las rodea. Con la hipótesis de una vida media para el protón de 10^{30} años, la intensidad de desintegración de una estrella típica de tamaño solar se cifrará en torno a los 10^{27} por año. Cada desintegración de un protón ocasiona una cascada de electrones, positrones, neutrinos y fotones de mucha energía. Todas las partículas hijas son absorbidas por la estrella; energía absorbida que mantiene caliente a la estrella.

Puede determinarse con precisión la temperatura de la estrella durante la época en que los protones se desintegran, igualando la intensidad de la radiación emitida por la estrella con la intensidad de la energía calorífica cedida por las desintegraciones. En este estado de equilibrio, la temperatura depende de la masa de la estrella, del área de la superficie, desde la que se emite el calor, y de la energía en reposo y vida media del protón. Esta temperatura de equilibrio es de 100 grados Kelvin para las estrellas muertas con mayor masa (que paradójicamente son las más pequeñas en tamaño) y de unos tres grados Kelvin para las mayores en tamaño y con menor masa.

Las estrellas alcanzarán su temperatura de equilibrio dentro de unos 10^{20} años. A partir de entonces, su temperatura permanecerá prácticamente constante hasta que casi todos sus protones se hayan desintegrado, hacia el año 10^{30}. Esta temperatura es fría, aunque no lo es en comparación con la temperatura de la radiación de fondo dejada por la explosión inicial. La temperatura de este fondo depende del desarrollo de la expansión del universo abierto. Si la densidad del universo es menor que la densidad crítica, la temperatura de la radiación de fondo descenderá hasta los 10^{-20} grados Kelvin hacia el año 10^{30}. Si la densidad es exactamente igual a la crítica, el universo se expansionará más despacio y la temperatura de la radiación de fondo descenderá hasta 10^{-13} grados Kelvin. En resumen, la temperatura será de 13 a 20 órdenes de magnitud inferior a las temperaturas de las estrellas muertas.

La desintegración de los protones también cambia la constitución de los gases interestelares que se evaporan antes del colapso de las galaxias. El positrón que aparece cuando se desintegra un protón dentro de una estrella encuentra pronto un electrón; las dos partículas se aniquilan mutuamente. Esta aniquilación genera más fotones y calienta la estrella. En el espacio intergaláctico, la densidad de materia es tan baja (y disminuye continuamente debido a la expansión del universo) que es muy poco probable que los electrones y positrones choquen. En realidad, hacia el año 10^{30}, un universo abierto con una densidad subcrítica se habrá expansionado hasta un tamaño 10^{20} veces el actual y la distancia media, en el espacio interestelar, entre un electrón y un positrón será del mismo orden de magnitud que el tamaño de nuestra galaxia. (Si la densidad del universo es igual a la densidad crítica, el universo se dilatará hasta un tamaño igual a 10^{13} el actual.) El medio interestelar será, pues, un gas muy enrarecido, formado por más o menos el uno por ciento de los electrones que actualmente hay y positrones creados por la desintegración de un uno por ciento del número de protones que hoy lo ocupan.

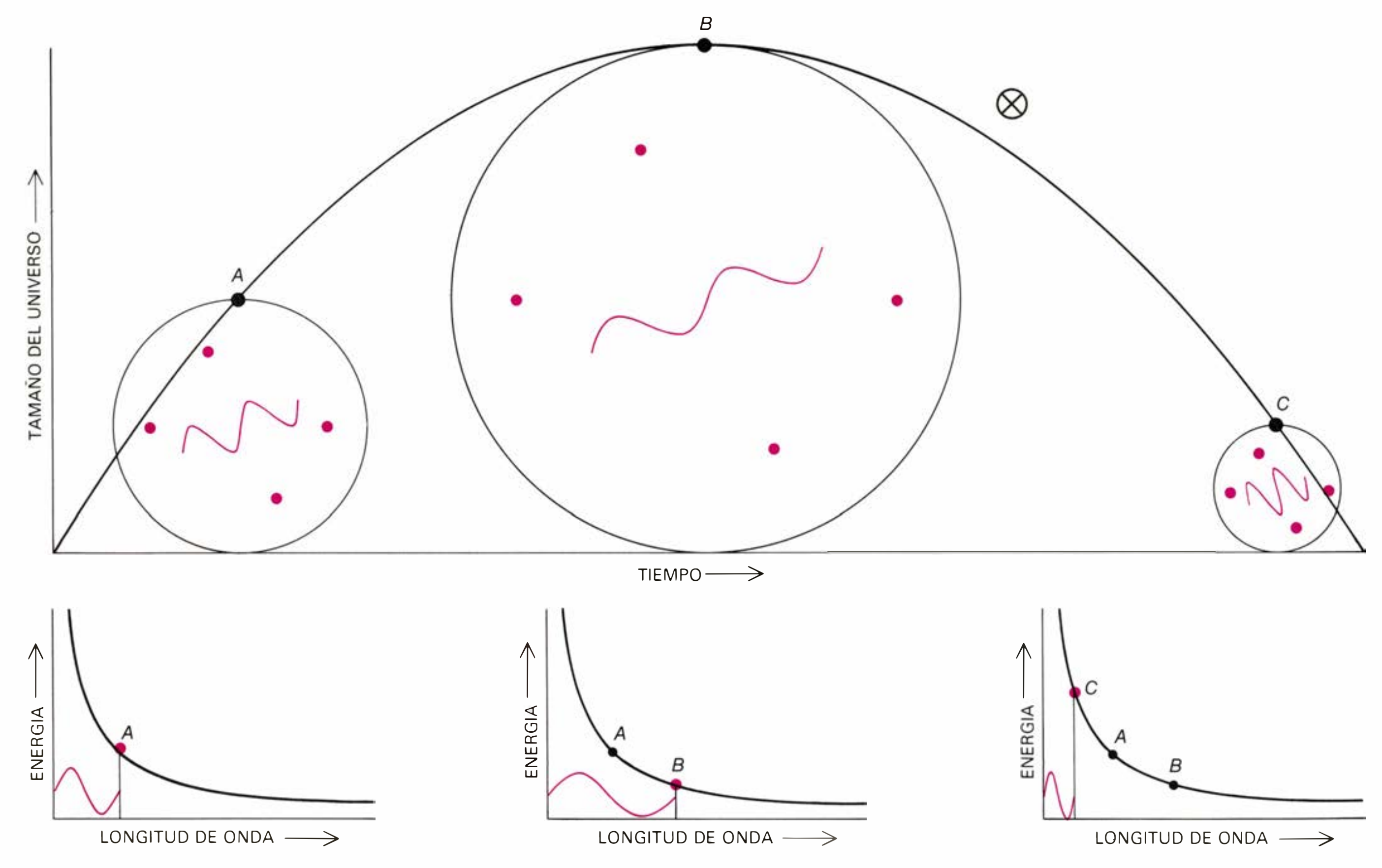
Los sucesos asociados con la desintegración del protón desaparecerán cuando el universo tenga 10^{32} años, o sea, cien veces más que la vida media del protón. Lo que quede entonces del universo será el gas enrarecido de electrones-positrones, los fotones y neutrinos que se emitieron en épocas anteriores y los agujeros negros de gran masa. Los fotones y neutrinos proceden de la explosión inicial, de los tiem-

pos cuando brillaban las estrellas, de la desintegración a lo largo de la historia de los protones y neutrones y de la desintegración final de las estrellas muertas. Los fotones y los neutrinos pierden energía; tanto ellos como los otros constituyentes van enrareciéndose, ya que la expansión continúa. En otros aspectos el universo con densidad subcrítica permanecerá estacionario hasta más o menos el año 10^{100}; este período, estacionario, se prolongará 10^{68} veces cuanto duran los sucesos que hasta ahora hemos descrito.

La sexta y última gran transición futura de un universo abierto asistirá a la desintegración de los agujeros negros. Nada puede escaparse de un agujero negro, según la interpretación más sencilla de la teoría gravitatoria de Einstein. Hay una frontera, llamada el horizonte de sucesos, en donde la velocidad necesaria para escapar es igual a la velocidad de la luz; ninguna partícula sita dentro del horizonte de sucesos alcanza la velocidad suficiente para atravesarlo. Sin embargo, en 1974, S. W. Hawking, de la Universidad de Cambridge, demostró que, gracias a un fenómeno de mecánica cuántica, un agujero negro puede entregar toda la energía asociada con su masa y desaparecer por consiguiente.

El fenómeno cuántico es consecuencia del principio de incertidumbre de Werner Heisenberg, que nos dice que, sólo con cierta probabilidad, podemos establecer que el momento o la posición de cualquier partícula están dentro de un determinado margen de valores. De una forma más precisa, el producto de la incertidumbre que tenemos sobre el momento de la partícula por la incertidumbre sobre la posición no es menor que una constante dada. En la física clásica (o sea en aquellas teorías físicas que no incluían fenómenos cuánticos), la partícula supera una barrera energética en el caso exclusivo de que reciba energía suficiente para sobrepasarla. El horizonte de sucesos supone una barrera energética absoluta para la física clásica: ninguna partícula puede obtener nunca suficiente energía para superarla. Merced al principio de incertidumbre, sin embargo, una partícula que inicialmente se halle en una región puede luego encontrarse en otra región, aun cuando su energía clásica sea mucho menor que la altura de la barrera de energía entre ambas regiones. De la partícula que atraviesa así una barrera de energía, esto es, sin adquirir la energía necesaria para superarla, se dice que ha sufrido el efecto túnel a través de la barrera. Hawking advirtió que los agujeros negros podían emitir masa y energía en virtud del efecto túnel de las partículas a través del horizonte de sucesos [véase "La mecánica cuántica de los agujeros negros", por S. W. Hawking, INVESTIGACIÓN Y CIENCIA, marzo, 1977].

Hawking demostró que la intensidad con la que un agujero negro emitía energía era inversamente proporcional al cuadrado de su masa. Esta intensidad, al principio, es baja; cuando la masa disminuye, la pérdida de energía se acelera. De ahí se deduce que todos los agujeros negros acaban por desaparecer o evaporarse; para los agujeros negros de gran masa, residuos del colapso de galaxias, el tiempo de evaporación es de unos 10^{100} años. La mayoría de los productos de la desintegración son fotones; las emisiones durante las últimas etapas de la evaporación son cada vez más energéticas. Así pues, tras 10^{100} años el universo estaría constituido por un gas extremadamente difuso de electrones, positrones y neutri-



7. LAS ENERGIAS DE LAS PARTICULAS LIBRES no se conservan en un universo cerrado. Debido a la expansión y contracción cósmica de este universo, los fenómenos que dependen de la distancia varían con el tiempo, del mismo modo que la distancia entre dos puntos de un balón crece o decrece según se hinche o deshinche el balón. Asimismo, la longitud de onda de un fotón crece durante la expansión y decrece duranta la contracción. La energía del fotón es inversamente proporcional a su longitud de onda; vale decir: la energía de un fotón emitido aumenta durante la expansión y disminuye durante la contracción. Cuando la longitud de onda del fotón sea más pequeña que la que tenía cuando se emitió habrá una ganancia neta de energía en el universo considerado en su totalidad. Esta energía adicional podría prolongar una nueva expansión subsiguiente. (Diseño realizado por Walken Graphics.)

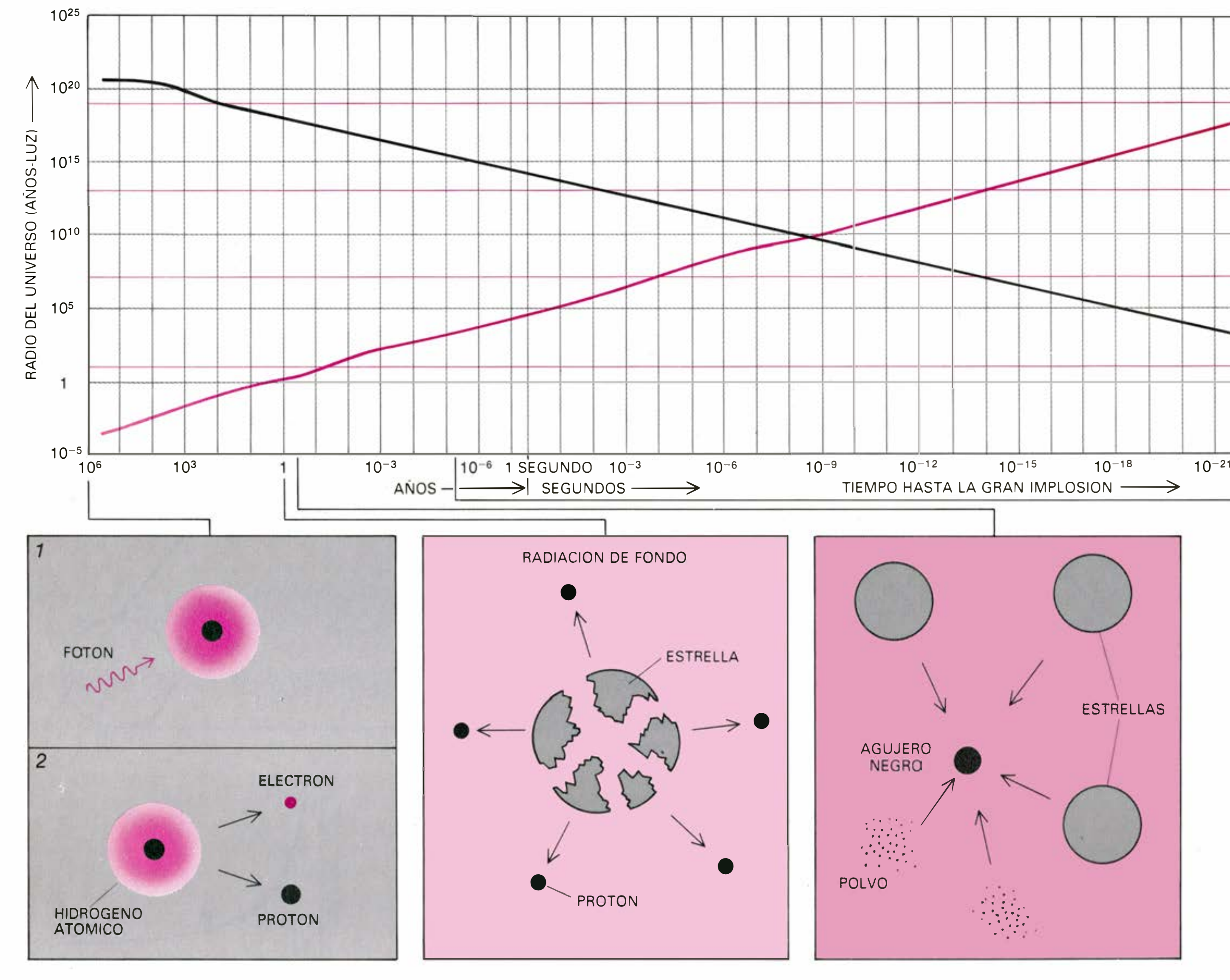
nos, de fotones de baja energía emitidos mucho antes de que los agujeros negros se desintegraran y de muchas esferas en expansión de fotones de alta energía emitidos por los agujeros negros al evaporarse.

Los modelos matemáticos de la estructura a gran escala de un universo abierto (de su macroestructura, podríamos decir) sugieren que el universo en que vivimos es un lugar muy especial. Para describirlo matemáticamente hemos de especificar seis magnitudes: tres que dan la intensidad de la expansión según tres direcciones del espacio y tres más que dan el valor del cambio en la intensidad de expansión (o su aceleración o deceleración) según esas tres direcciones. Si, en un instante considerado, establecemos estas seis cantidades, así como la distribución de energía en ese mismo instante, quedarán determinados para cualquier tiempo posterior las seis magnitudes y la distribución de energía.

En 1973, Hawking y C. B. Collins, también de Cambridge, demostraron que, para casi todos los valores posibles de estas cantidades en una primera etapa de la gran explosión, se desarrolla una anisotropía que aumenta con la expansión. La separación media de galaxias según un eje del espacio debería diferir, cada vez más, de la separación media a lo largo de otros ejes por causa de estas discrepancias en las velocidades de expansión. Este resultado sugiere que el universo real es un universo muy poco probable. Como hemos dicho antes, la radiación cósmica en microondas no varía con la dirección, con sólo un posible error de una parte sobre 10.000. A una escala cósmica, también la materia del universo está distribuida muy uniformemente. Uno de los problemas por resolver en cosmología es el de explicar por qué el universo está tan cerca de la isotropía.

Otra forma alternativa de abordar el problema: escoger valores iniciales de las seis magnitudes, de suerte que la densidad energética de masa y radiación sea exactamente igual a la densidad crítica. Si, a lo largo de su historia, este universo estuviese dominado por la materia y no por la radiación, no prosperaría ninguna anisotropía inicial. Sin embargo, la desintegración del protón puede conducir a un universo dominado por la radiación.

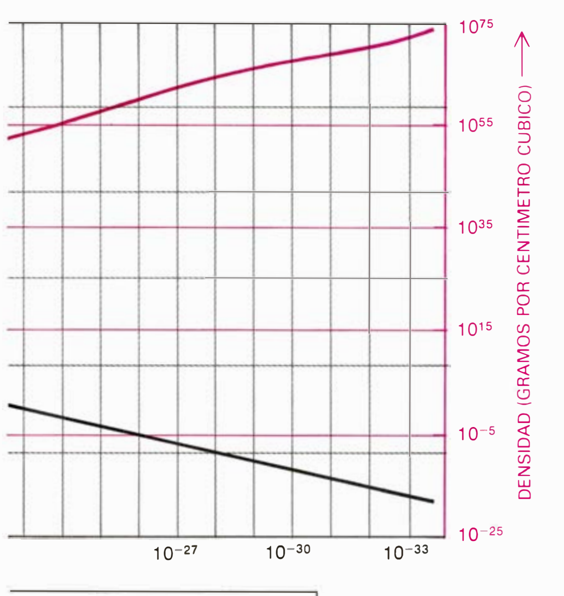
John D. Barrow y Frank J. Tipler, de la Universidad de California en Berke-

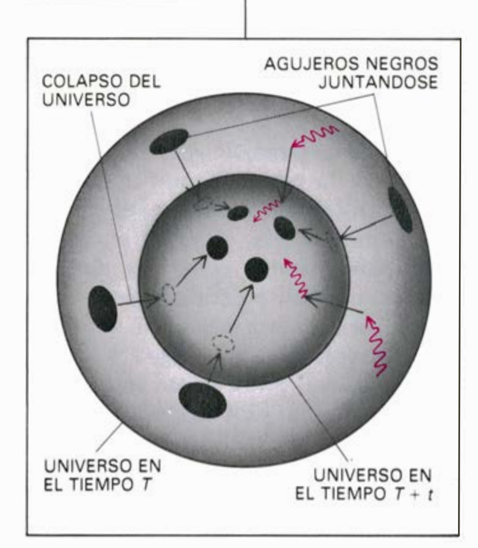


8. ETAPAS FINALES del colapso de un universo cerrado. Remedaría el curso inverso de las etapas de la expansión, si no fuera por los efectos provocados por los agujeros negros. Cuando falten unos 20.000 millones de años para la gran implosión ("the big crunch"), o momento del colapso gravitatorio completo del universo cerrado, el universo se hallará tan contraído que su densidad de enegía será la misma que tenemos ahora. Al continuar la contracción, los fotones irán adquiriendo mayor energía y el universo se calentará; a unos millones de años de la gran implosión, los fotones provocarán la disociación de los átomos de hidrógeno interestelares en sus constituyentes: electrones y protones. Un año antes de la implosión, la temperatura en el exterior de las estrellas será mayor que la interior y las estrellas empezarán a desintegrarse. Hacia la misma época, los agujeros negros de gran masa iniciarán la absorción de la materia estelar que quede, así como de cualquier otra clase de materia o radiación. Unos tres minutos antes de la implosición, los agujeros

ley, han demostrado que en un universo con la densidad crítica los electrones y positrones resultantes de la desintegración de los protones iniciarán la formación de pares ligados cuando hayan transcurrido más de 10^{70} años. Cada par ligado es, en realidad, un átomo, pero las órbitas en que el electrón y el positrón giran uno alrededor del otro pertenecen a un estado tan fuertemente excitado que la región del espacio que abarca el par es mayor que el universo que nosotros observamos. Este espacio abarcado va aumentando a medida que va tardando en formarse el par ligado.

Conforme transcurre el tiempo, el electrón y el positrón describen espirales que se cierran y al final se aniquilan mutuamente produciendo fotones de alta energía. La longitud de onda de estos fotones tiene un corrimiento hacia el rojo del espectro a medida que el universo continúa su expansión. En este modelo de universo, descrito por Barrow y Tipler, el universo acaba dominado por la radiación y crece la anisotropía. Sin embargo, en una investigación posterior, Don N. Page y M. Randall McKee, de la Universidad estatal de Pennsylvania, demuestran que en ciertas condiciones los pares electrón-positrón continuarían dominando la energía del universo, aunque su número decreciera. El corrimiento hacia el rojo de la energía del fotón disminuye su energía; el tiempo, cada vez mayor, que se requiere para aniquilar los pares formados posteriormente provoca una disminución de la intensidad energética de la emisión de fotones. Por tanto, la anisotropía, característica de un universo dominado por la radiación, quizá no llegara a desarrollarse.





negros de gran masa empezarán a juntarse; aquí esta unión aparece dibujada en la superficie de una esfera que se contrae, lo que corresponde a la contracción del espacio. La gráfica de la variación de la densidad y el radio del universo respecto del tiempo; los cálculos se basan en las ecuaciones de Einstein y otros principios descritos en el texto.

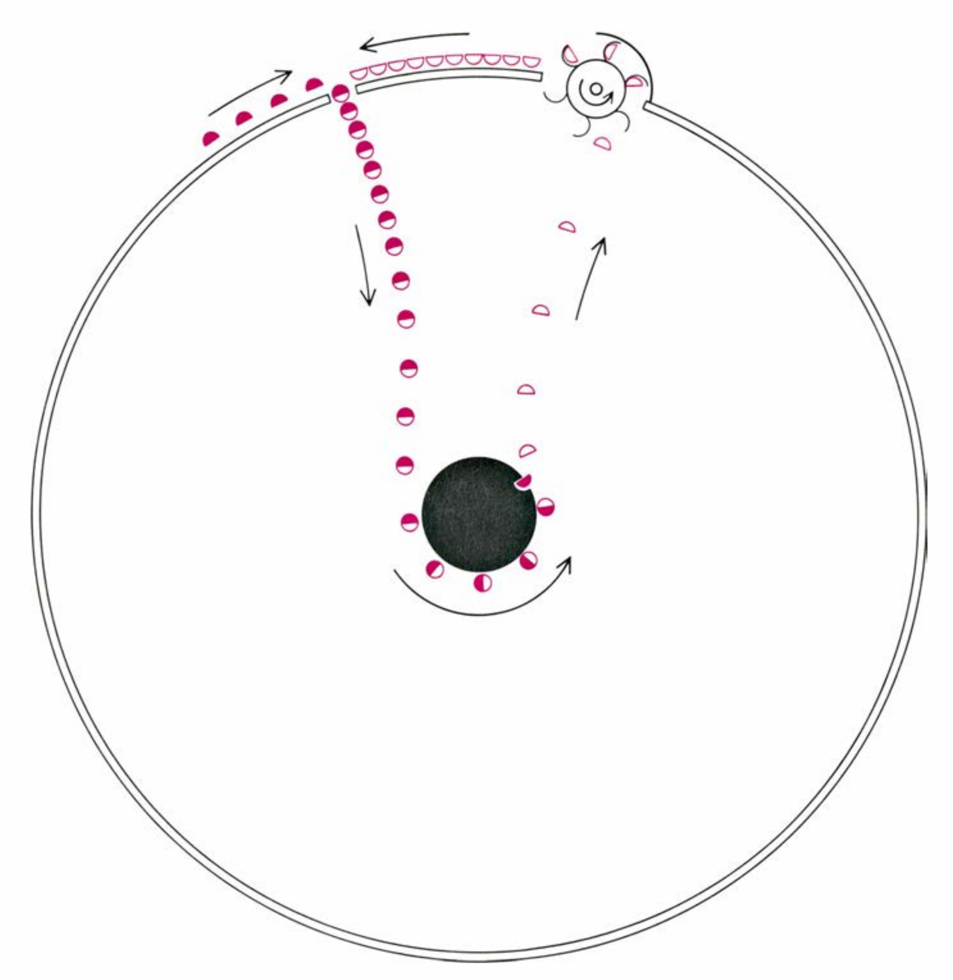
Todos los razonamientos anteriores corresponden a un universo abierto. Consideremos ahora el futuro del universo en el supuesto de que haya suficiente materia no luminosa para que la gravedad detenga la expansión y provoque una contracción. Cuanto más cerca de la densidad crítica esté la densidad media, tanto más durará el universo cerrado. No sabemos de ninguna causa que justifique que la densidad media no esté lo suficientemente cerca de la densidad crítica, de suerte que el universo pueda sobrevivir hasta que la mayoría de los protones se hayan desintegrado. Lo más probable, pues, es que el universo, en el punto de máxima expansión, estuviese formado por estrellas muertas, agujeros negros de gran masa, como restos de las galaxias colapsadas, y fotones y neutrinos de baja energía, lo mismo que si se tratara de un universo abierto.

El universo cerrado ofrece una peculiaridad curiosa: aunque la energía se conserve localmente, no se conserva la masa o energía total del universo. Para cierto tamaño del universo, la energía total durante la contracción es mayor que la existente durante la expansión. Consideremos un fotón emitido por el Sol hacia el espacio intergaláctico. Durante la emisión, se cumple la ley de conservación, ya que la energía con que sale el fotón se compensa exactamente con un pequeño decrecimiento de la masa del Sol. A medida que el universo se expande, la longitud de onda del fotón aumenta proporcionalmente, y así su energía disminuye. Al contraerse, la longitud de onda del fotón se contrae también, por cuya razón aumenta su energía. La longitud de onda del fotón acaba por ser más corta que lo era en el instante de emisión: el fotón habrá ganado energía sin que ocurra, en compensación, ninguna pérdida de masa o energía en otra región. De donde se deduce que la evolución del universo es más caliente en la contracción que durante la expansión. Los fotones que contribuyen más a la energía suplementaria son los emitidos cuando el universo se halla cerca de su tamaño máximo, ya que estos fotones sufrirán la máxima contracción.

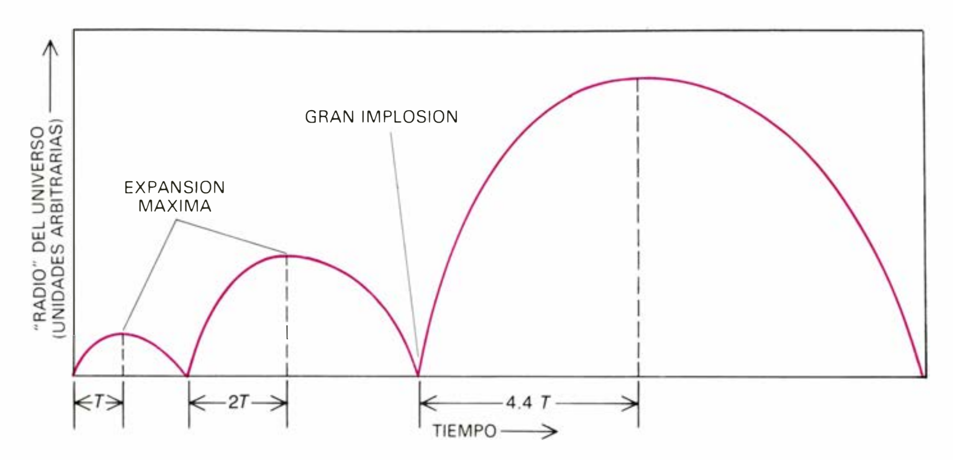
Las etapas principales que ocurren durante la fase de expansión de un universo cerrado siguen la misma secuencia que las etapas de un universo abierto. (Sin colapso alguno para un universo abierto, por supuesto.) Varios investigadores, entre ellos Martin J. Rees, de Cambridge, han estudiado el colapso. Cuando los fotones ganan energía a lo largo de la contracción, calientan las estrellas muertas, provocando el inicio de una rápida combustión, su explosión o su evaporación. Las partículas que quedan continuarían siguiendo los pasos de la expansión del universo, si no fuera por los efectos de los agujeros negros. Al aumentar la densidad, los agujeros negros tragan más materia; cuando no chocan, se unen. Cabe suponer que en un universo, con un agujero negro de gran masa por galaxia, éste engulle las estrellas muertas al poco de romperse éstas y en cuanto empiece a evaporarse materia de las mismas. Al cabo, todos los agujeros negros se funden en un solo y gran agujero negro tan extenso como el universo.

La física teórica no sabe explicar todavía, en todos sus pormenores, el colapso de un agujero negro. Le faltan conocimientos para extrapolar las ecuaciones que gobiernan la macroestructura del universo en un estado de densidad infinita. Pero cabría la posibilidad de que antes de que la densidad se haga infinita se presentara un mecanismo desconocido que provocara un "rebote" del universo y que éste empezara de nuevo a expansionarse.

Si el universo rebotara, podría permanecer cerrado y, por tanto, ser cíclico. La energía ganada por los fotones durante cada período de contracción podría conservarse en el rebote; de este modo, en cada ciclo sucesivo el universo sería, para una temperatura dada, mayor que en el ciclo anterior y le costaría también más tiempo llegar a su tamaño máximo. Robert H. Dicke y P. J. E. Peebles, de la Universidad de Princeton, y el cuarto de los que firmamos este artículo han calculado que, con un



9. **LA CRISIS ENERGETICA A NIVEL COSMICO puede retrasarse gracias a la existencia de los agujeros negros procedentes del colapso de las galaxias. En teoría, podemos extraer energía del campo gravitatorio de un agujero negro. Entra, pues, dentro de lo imaginable que una civilización avanzada subsistiera hasta el año 10^{100}, o sea, hasta que los agujeros negros se evaporaran. De todos los conocidos, el esquema más eficaz para extraer energía de un agujero negro fue sugerido por Charles W. Misner, de la Universidad de Maryland, Kip S. Thorne, del Instituto de Tecnología de California, y John Archibald Wheeler, de la Universidd de Texas en Austin. Se construye una esfera hueca rígida, que envuelva a un agujero negro que gira, y se envía un satélite con materiales de desecho de la civilización hacia el horizonte de sucesos, la frontera de la región de donde nada puede escaparse. Cerca de dicho horizonte de sucesos se lanza la basura sobre el ecuador del agujero negro. La energía potencial gravitatoria del material desechado y la energía de rotación del agujero negro disminuyen. Puesto que debe conservarse la energía total del satélite más la del agujero negro, la energía del satélite ha de aumentar en una cantidad igual a las dos pérdidas citadas. La energía ganada por el satélite puede convertirse en otra forma aprovechable.**



10. **POSIBILIDAD DE UN UNIVERSO CICLICO. Se acepta si la densidad de materia es suficiente para cerrarlo y si procesos físicos hasta ahora desconocidos le permiten "rebotar" y expansionarse otra vez. Como en la fase de contracción se gana energía, cada nueva expansión dura más tiempo que la del ciclo precedente. El período sería el doble en el ciclo siguiente del universo, partiendo de que la expansión actual durara 100 veces la edad que tiene hoy. En los ciclos sucesivos la duración iría prolongándose.**

universo cíclico, en el ciclo siguiente la fase de expansión duraría unas dos veces la actual. Para ciclos anteriores, el factor de expansión sería más pequeño; según nuestras estimaciones, el presente universo está, como máximo, a cien rebotes del ciclo que duró lo suficiente para conseguir que se formara una sola generación de estrellas. La principal dificultad que se le plantea al modelo de rebotes es la de explicar de qué manera puede regularizarse un universo de agujeros negros que se juntan, extraordinariamente inhomogéneo y localmente anisótropo. Para una materia así, no hay razonamiento que justifique la posibilidad de un rebote.

Si el universo es cíclico y si las reacciones provocadas por las altas temperaturas existentes al principio de cada ciclo causan el exceso, que hoy observamos, de partículas nucleares respecto de antipartículas, podríamos llegar a una clara conclusión: la evolución del universo desde el inicio de cada rebote hasta la formación y disolución de galaxias seguiría la misma secuencia de etapas en cada ciclo. Sin embargo, cuando los ciclos duren más, los fenómenos que hemos descrito para las últimas etapas del universo abierto cobrarán una importancia mayor. Por ejemplo, en el ciclo en que la expansión dure, por primera vez, más de la diezmilmillonésima parte de la vida media del protón (esto es, más de 10^{20} años) la energía ganada a lo largo del período de contracción se debería fundamentalmente a los fotones emitidos en la desintegración de los protones. Los ciclos siguientes podrían entonces alargarse en un factor 1000, y no en un factor de alrededor de dos. Podrían conseguirse factores de expansión todavía mayores, del orden de 10^{12}, a partir de los efectos de la unión de agujeros negros.

Desde un punto de vista antropológico, la pregunta más apremiante se cifraría en torno al futuro de la vida y de la inteligencia. ¿Pueden las formas del pensamiento, como las que hay ahora, persistir indefinidamente en un universo cada vez más hostil a la vida? Varios cosmólogos, Dyson y Steven C. Frautschi, entre ellos, han comenzado a abordar problemas de índole parecida al siguiente: ¿cuáles son las necesidades energéticas que requiere el mantenimiento indefinido de la vida y la comunicación entre los miembros de una sociedad diseminados a lo largo de regiones, cada vez más extensas, del espacio?

Dyson supone que la encarnación de la vida y la conciencia no estaría limitada a las células y al ADN. Antes bien,

sostiene que el hecho esencial de la conciencia es la complejidad de la estructura, que podría alcanzarse con cualquiera material idóneo. De ahí que admita que la idea de un computador o de una acumulación que puedan sentir sensaciones no deba rechazarse como posibilidades carentes de coherencia filosófica.

Enunciadas, pues, estas hipótesis, los cambios causados en el entorno por la muerte y enfriamiento de las estrellas y su evaporación de las galaxias no deberían resultar insuperables para un sistema que podría considerarse vivo e inteligente. Por ejemplo, podría en principio extraerse energía del campo gravitatorio de un agujero negro de gran masa. La desintegración de protones y neutrones aportaría un cambio fundamental, ya que parece improbable (aunque tal vez no sea imposible) que la inteligencia se base en un fondo tenue de electrones y positrones. Además, si el universo es cerrado, las condiciones precisas para la vida se darían sólo en ciertos períodos de cada ciclo.

En un universo abierto, la dificultad principal para la vida es de otro orden completamente distinto. Con la evaporación de los agujeros negros, habrá pérdida de energía cósmica: al proseguir la expansión del universo, las partículas y fotones que quedan cederán energía. Para las formas vivientes les resultará insostenible una tasa constante de consumo de energía. Como alternativa, Dyson propone la intercalación de períodos de hibernación cada vez más prolongados, en los que no se consumiría energía, entre períodos de consumo. Por tanto, este universo encierra potencial suficiente para civilizaciones de larguísima duración.

La cosmología contemporánea arranca de la teoría de la gran explosión. Nuestro tratamiento cuantitativo de la evolución de la gran explosión en un futuro lejano se funda en las nuevas teorías que describen la física de las partículas elementales. Pero, en espíritu, estas ideas las expresó ya en el siglo XI el astrónomo y poeta persa Omar Jayyam, cuyo poema "El Rubáiyát" tradujo al inglés Edward FitzGerald. Así leemos en la cuarteta LIII:

> Con la primera arcilla de la Tierra elaboraron la masa de la que se hizo el último hombre,
> Luego sembraron la simiente de la última cosecha:
> Sí: la primera mañana de la Creación escribió
> Lo que leerá el último amanecer del Día del Juicio.

La memoria de las aves que acaparan alimento

Las aves que esconden semillas para luego recuperarlas sobresalen por su memoria espacial. Algunas recuerdan, meses enteros, la localización de miles de escondrijos

Sara J. Shettleworth

En el invierno inglés, un comedero de pájaros abarrotado de cacahuetes es todo el aparato que se necesita para poner de manifiesto la memoria tenaz de muchas aves. Facultad que parece sobrepasar, con mucho, la que cualquier otro animal haya mostrado nunca en el laboratorio. El comedero atrae a muchos pájaros: carboneros comunes, carboneros palustres y herrerillos comunes, pájaros pequeños y bulliciosos de la familia Páridos, entre otros. Carboneros y herrerillos comunes se congregan en el comedero, y de inmediato empiezan a comer. Sólo interrumpen su banquete para alejar a los competidores. El carbonero palustre, en cambio, llega en rápido vuelo, toma un cacahuete y sale disparado. No tardará mucho en volver por otro. Los almacena en las inmediaciones, cada uno en un lugar distinto, hasta que el comedero queda vacío. Y comienza entonces a recobrar el alimento escondido.

En el sudoeste americano, el cascanueces de Clark, pariente de arrendajos y cuervos, muestra un comportamiento bastante similar. A finales de verano recolecta las semillas de pinos piñoneros; una y otra vez llena su bolsa sublingual, un receptáculo situado bajo la lengua, y vuela varios kilómetros para enterrar los piñones. A menudo los lugares de enterramiento se hallan en pendientes desnudas orientadas al sur, donde la nieve no alcanzará, en su momento, mucha altura. Un cascanueces entierra hasta 33.000 piñones en escondites de cuatro a cinco semillas cada uno. Durante el invierno, retornará y excavará sus miles de escondrijos.

¿De qué se valen estas aves para encontrar sus almacenes? ¿Acaso el carbonero o el cascanueces recuerdan dónde han guardado cada cacahuete, la ubicación de cada escondrijo de piñones? Hasta hace poco, los estudiosos del comportamiento de acaparamiento en la naturaleza dudaban que las aves se apoyaran en la memoria. De entrada, el ave necesitaría una memoria extensa que le permitiera recordar los lugares de cientos o miles de objetos almacenados uno a uno. Asimismo, la memoria tendría que ser duradera. El mismo acaparador a corto plazo, el carbonero palustre por ejemplo, no recupera su despensa hasta pasadas horas o días desde que la depositó; el acaparador a largo plazo, así el cascanueces de Clark, tardará meses en acudir a sus escondrijos, si es que no espera hasta la primavera, estación en que alimenta con piñones a sus pollos. Además, los acúmulos podrían localizarse, en principio, sin ayuda de la memoria. Un acaparador podría almacenar alimento únicamente en determinado tipo de lugares, como laderas orientadas al sur o agujeros de la corteza de árboles. Para recuperar sus provisiones sólo precisaría buscar en lugares de las características adecuadas. En este caso, podría emplear la memoria sólo para recordar en qué zona puso a recaudo el alimento, en vez de recurrir a ella para cada lugar en particular. En esa zona podría buscar por tanteo, o guiarse por señales, como el olor.

Sin embargo, los psicólogos que estudian el aprendizaje y la memoria en los animales están cada vez más convencidos de que determinadas especies poseen especializaciones adaptativas que las hacen particularmente eficaces a la hora de aprender y recordar datos que les son importantes. Entre los ejemplos bien conocidos se halla la capacidad de muchas aves para aprender los cantos de su especie, la habilidad de ratas y otros animales para recordar posiciones en el espacio, la capacidad de las abejas para rememorar el emplazamiento de las flores y la aptitud de muchos animales para evitar alimentos nocivos. De ser cierto que las aves que almacenan comida recuerdan gran número de lugares de acopio durante períodos prolongados, su memoria constituiría otro ejemplo de especialización adaptativa, especialización que les permitiría recuperar con eficacia sus provisiones. El ave que recordara dónde escondió su alimento cometería menos errores a la hora de recuperarlo y gastaría menos tiempo y energía que las que buscaran al azar. Estudios recientes indican que al menos algunas aves que almacenan alimento recuerdan bastante bien los emplazamientos de sus provisiones.

John R. Krebs y sus colegas, de la Universidad de Oxford, han reunido estos últimos años buena parte de las pruebas más detalladas de que las aves acaparadoras de alimento recuerdan sus escondrijos. Empezaron buscando pruebas que confirmaran que los carboneros palustres, en su ambiente natural, recuperaban sus provisiones. Con este fin entrenaron a carboneros palustres en la acción de dirigirse a comederos automáticos que suministraban pipas de girasol, situados en sus territorios de Wytham Wood, cerca de Oxford. Richard J. Cowie, Krebs y David F. Sherry recubrieron entonces las cáscaras de las pipas con una sustancia radiactiva inofensiva para los pájaros. (Estos descascaran las pipas antes de comérselas.) La radiactidad permitió a los investigadores, equipados con un contador de centelleo portátil, encontrar las semillas que las aves habían escondido y comprobar cada pocas horas si seguían en su sitio.

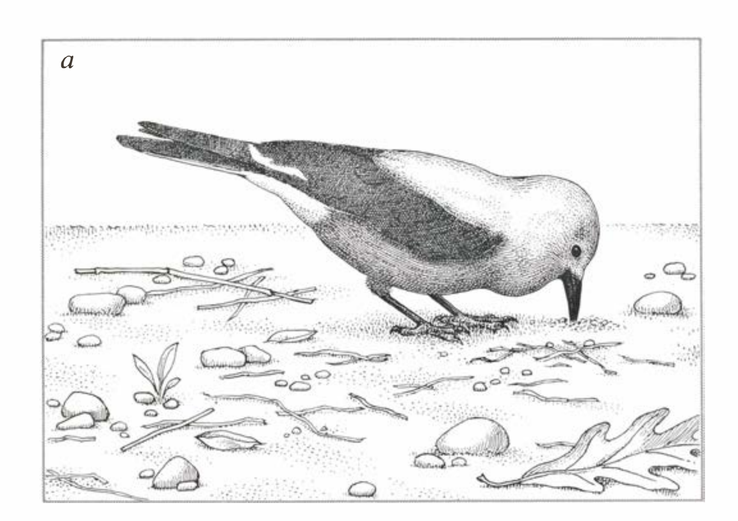
Se corría el riesgo ineludible de que otras aves o los mismos roedores, como ratones y topillos, llegaran antes a las pepitas. No obstante, razonaron Co-

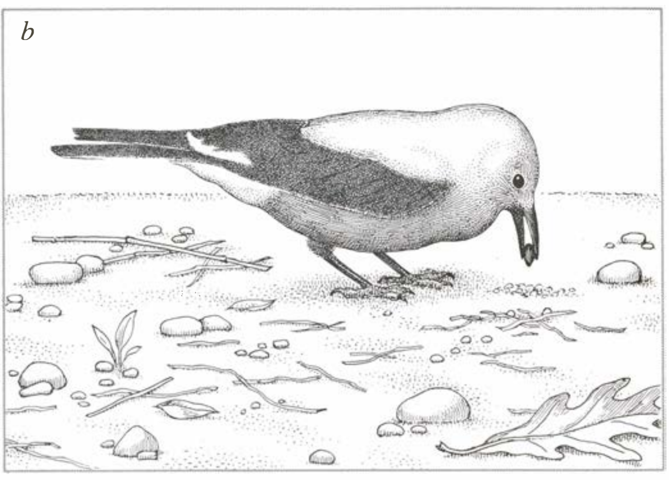
wie, Krebs y Sherry, si los carboneros palustres empleaban la memoria para recuperar sus provisiones, las semillas desaparecerían antes de los lugares de almacenamiento de los pájaros que de los falsos emplazamientos que los investigadores establecieron. En efecto, así ocurrió. Las pipas que los pájaros acumularon se agotaron antes que las que los investigadores colocaron en lugares similares, todos, respectivamente, a un metro de algún acúmulo reunido por los pájaros. En Wytham Wood, en invierno, los acúmulos naturales desaparecían por lo general al cabo de un día o dos. El último de los acúmulos falsos desapareció muy poco después, lo que sugiere que la depredación de estos almacenes era habitual.

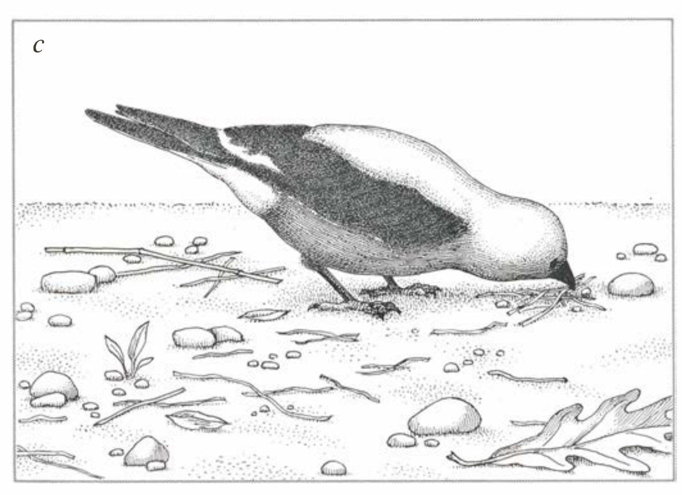
Frente a los animales que roban sus almacenes, los carboneros palustres parecen adoptar la táctica defensiva de dispersar las semillas, y no enterrarlas en grupo. Por término medio, los acúmulos que los carboneros palustres establecieron en Wytham Wood distaban siete metros unos de otros. Sherry, Mark Avery y Allen Stevens dispusieron falsos almacenes de pipas de girasol y encontraron que las pepitas permanecían más tiempo en su lugar cuanto más alejadas estaban unas de otras, hasta una distancia de unos siete metros. Los acúmulos separados más allá de esta distancia no ofrecían mayor seguridad que los situados a ese intervalo. En Wytham Wood, al menos, los carboneros palustres parecen haber dado con el espaciamiento óptimo.

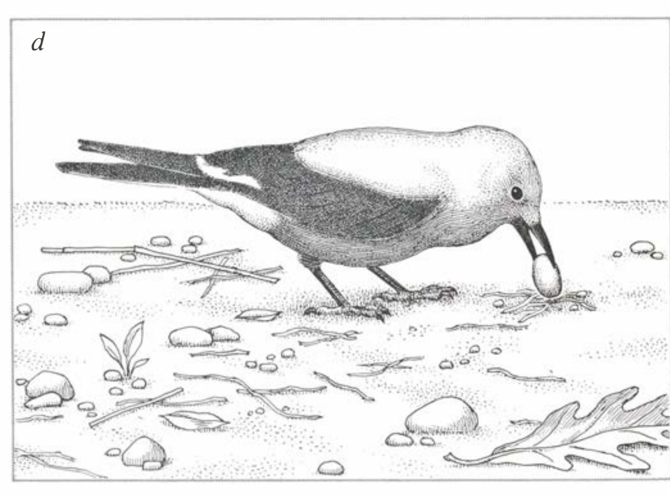
Confirmado ya que los carboneros palustres en libertad recuperan sus acúmulos, Krebs y sus colegas comenzaron los trabajos de laboratorio. En cautividad, los carboneros palustres a los que se suministra un cuenco de semillas y lugares de almacenamiento adecuado desarrollan fácilmente el comportamiento de acumulación. Sherry, Krebs y Cowie ofrecían a los carboneros palustres cajones de musgo de un metro de lado donde almacenar pipas de girasol. Los pájaros parecían recordar dónde habían enterrado las pepitas a las 24 horas. Sin embargo, en esos experimentos, el comportamiento de los pájaros se registraba únicamente en forma de visitas a zonas relativamente amplias de la gaveta. Además, para asegurarse de que se valían únicamente de la memoria se retiraban del cajón las semillas escondidas antes de devolver el pájaro a la estancia. Resultaba difícil, por tanto, decir con qué precisión los páridos localizaban sus lugares de almacenamiento. No más sencillo era identificar los errores que podían estar cometiendo.

Una beca Guggenheim me permitió pasar un año sabático en Oxford, donde mi interés por las especializaciones adaptativas del aprendizaje y la memoria me llevó a los estudios sobre el carbonero palustre. Krebs y yo diseñamos algunos experimentos para analizar con





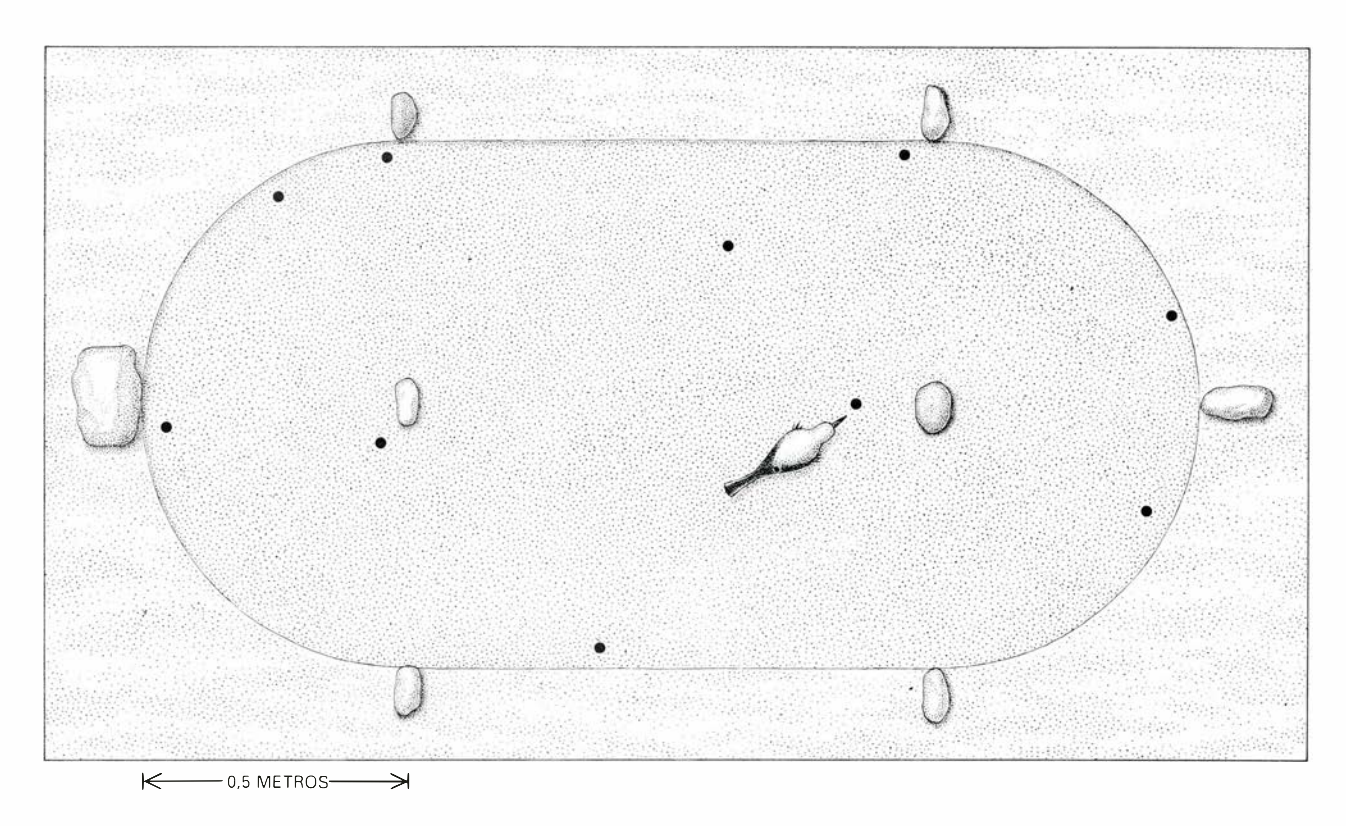


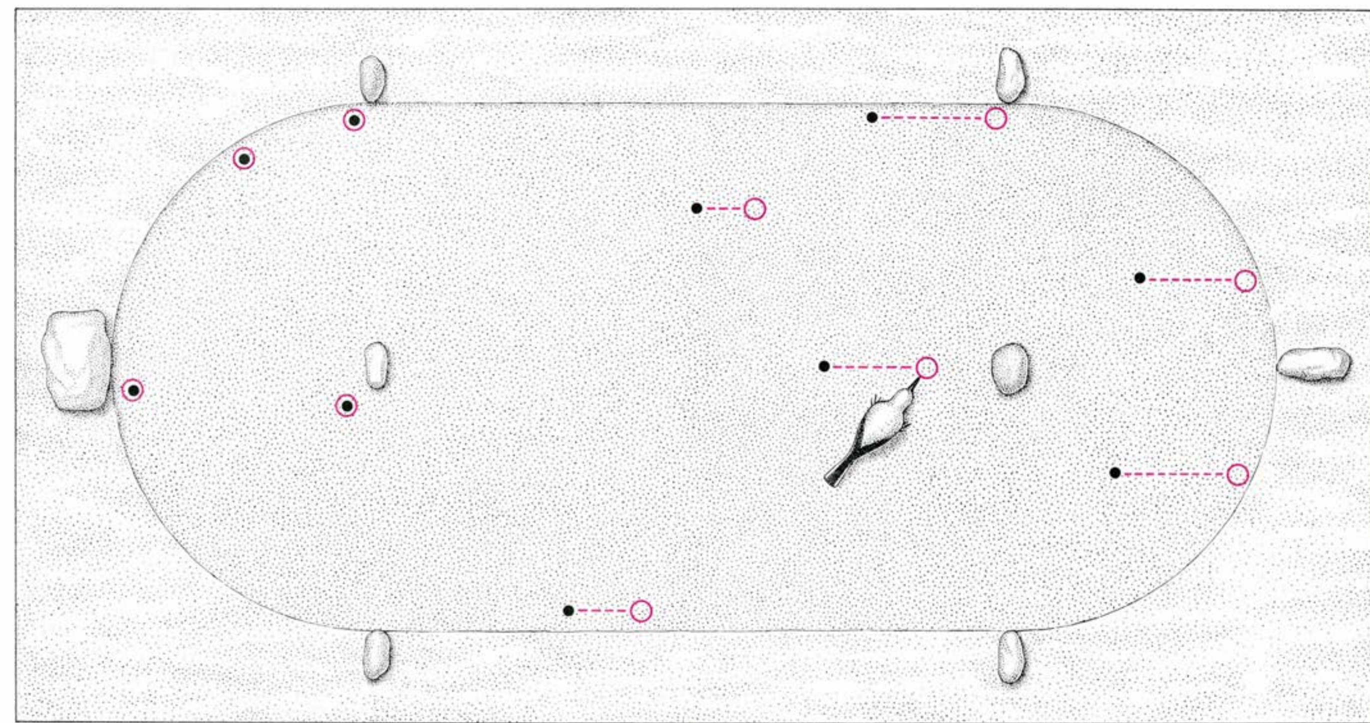


1. ACAPARAMIENTO DE ALIMENTO por un cascanueces de Clark, según se muestra en una secuencia de dibujos basada en observaciones de Stephen B. Vander Wall, de la Universidad estatal de Utah. El cascanueces remueve el suelo con su pico (*a*); luego deja una semilla de conífera (*b*), rastrilla suelo y hierbas sobre el piñón (*c*) y coloca un guijarro sobre el escondrijo (*d*). Aparentemente, el guijarro sólo sirve para enmascarar el lugar hasta que la intemperie barra toda traza del escondrijo. El proceso dura de 10 a 20 segundos. En la naturaleza, este cascanueces acapara hasta 33.000 piñones en miles de escondrijos ampliamente distribuidos, que tienen, por término medio, cuatro o cinco piñones cada uno. El ave acapara semillas de septiembre a noviembre, y recurre a la memoria para encontrar los escondrijos a lo largo del invierno, primavera y verano siguientes. La ilustración es de Tom Prentiss.

más detalle la memoria aparente de estos páridos. Nos interesaba en particular averiguar si los últimos objetos almacenados por los pájaros eran los que se recuperaban primero. Este "efecto de inmediatez" aparece con frecuencia en las pruebas de memoria, y hay una buena razón para esperarlo en el caso del acaparamiento de alimento: si los acúmulos se pierden por acción de los depredadores a un ritmo constante, los acúmulos más recientes son los que con mayor probabilidad seguirán disponibles para los acaparadores en un momento dado. Por tanto, dirigirse en primer lugar a ellos es la mejor opción.

En una sala amplia, Krebs y yo dispusimos secciones de ramas de árbol. En cada sección perforamos varios agujeros de unas medidas idóneas para





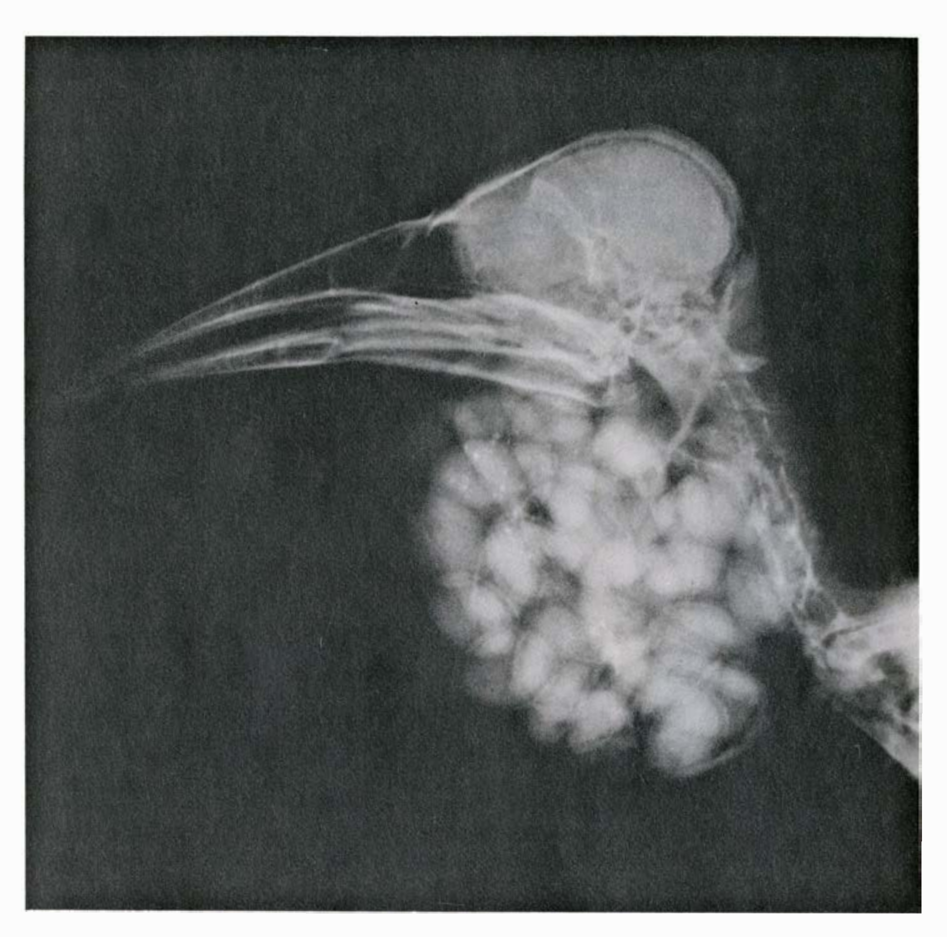
2. PRUEBA DEL EMPLEO DE LA MEMORIA por los cascanueces de Clark. Se obtuvo en experimentos realizados por Vander Wall. Se permitía que un cascanueces escondiera piñones (*puntos negros*) en el interior de una zona oval del suelo arenoso de un aviario; se dispusieron cuatro piedras a cada extremo del óvalo (*ilustración superior*). Se alargó luego el óvalo 20 centímetros y las piedras de la derecha se desplazaron en consecuencia (*ilustración inferior*). Dos días después se permitió que el ave buscara sus escondrijos. El cascanueces excavó con éxito los escondrijos de la izquierda (*círculos de color*). Sin embargo, a la derecha cometió errores de unos 20 centímetros. Resulta evidente que fiaba en su memoria de las posiciones de las piedras cercanas. Los errores cometidos en la parte media sugieren que el ave estimó la posición de algunos escondrijos con referencia a ambos grupos de hitos.

almacenar cañamones; en total había unos 100 agujeros, cubiertos todos ellos por un halda de tela que los pájaros debían levantar para almacenar una semilla o para recuperarla.

Desarrollamos en la estancia una serie de experimentos. En cada prueba de las que componían el primero permitíamos que un carbonero palustre almacenara 12 semillas, cobradas de un cuenco situado sobre el suelo, en el centro de la habitación. Los agujeros sólo podían alojar una semilla, de modo que cada cañamón se almacenaba en un agujero diferente. Después de que escondiera las semillas, manteníamos el ave fuera de la habitación unas dos horas y media. Retirábamos el cuenco de cañamones y devolvíamos el pájaro a la sala para que buscara las semillas que había escondido. Si el ave hubiera investigado al azar entre los 100 agujeros, habría escudriñado una media de unos ocho agujeros para encontrar una semilla. La eficacia de los páridos se demostró muy superior. Por término medio, cada individuo cometía un par de errores por semilla. A veces, al comienzo de una prueba de recuperación, un pájaro se dirigía sucesivamente a tres o cuatro cañamones sin buscar agujeros vacíos.

Quizás en lugar de emplear la memoria los carboneros olfatearan las semillas, o las detectaran de alguna manera en la que no habíamos pensado. El olfato parecía improbable, pues se trata de un sentido de escasa potencia en las aves. No obstante, para poner a prueba dichas alternativas, dejamos que un carbonero palustre almacenara 12 semillas y las trasladamos luego a otros agujeros (utilizados en pruebas anteriores por el mismo párido). En estas condiciones, las aves cometían unos seis errores por semilla. Después de buscar en más agujeros de los acostumbrados, acababan encontrando unos dos tercios de los cañamones.

Los resultados indican que los carboneros palustres no detectan guiados por el olfato, o por alguna señal similar, los cañamones almacenados. Por otra parte, la actuación de los carboneros en el primer experimento no significa necesariamente que recuerden el escondrijo preciso de cada semilla. En cierta medida, cada pájaro del laboratorio llegó a tener agujeros favoritos, en los que almacenaba semillas con mayor frecuencia. Cada pájaro tendía a inspeccionar sus agujeros favoritos cuando iba en busca de cañamones almacenados. La recuperación de éstos a partir de tales agujeros podía deberse únicamente a que las aves estaban acostumbradas a visitar esos lugares, para guardar o recuperar las semillas. Esta tendencia explica probablemente por qué la actuación de los páridos era mejor que la fundada en el mero azar cuando cambiábamos de lugar los cañamones.

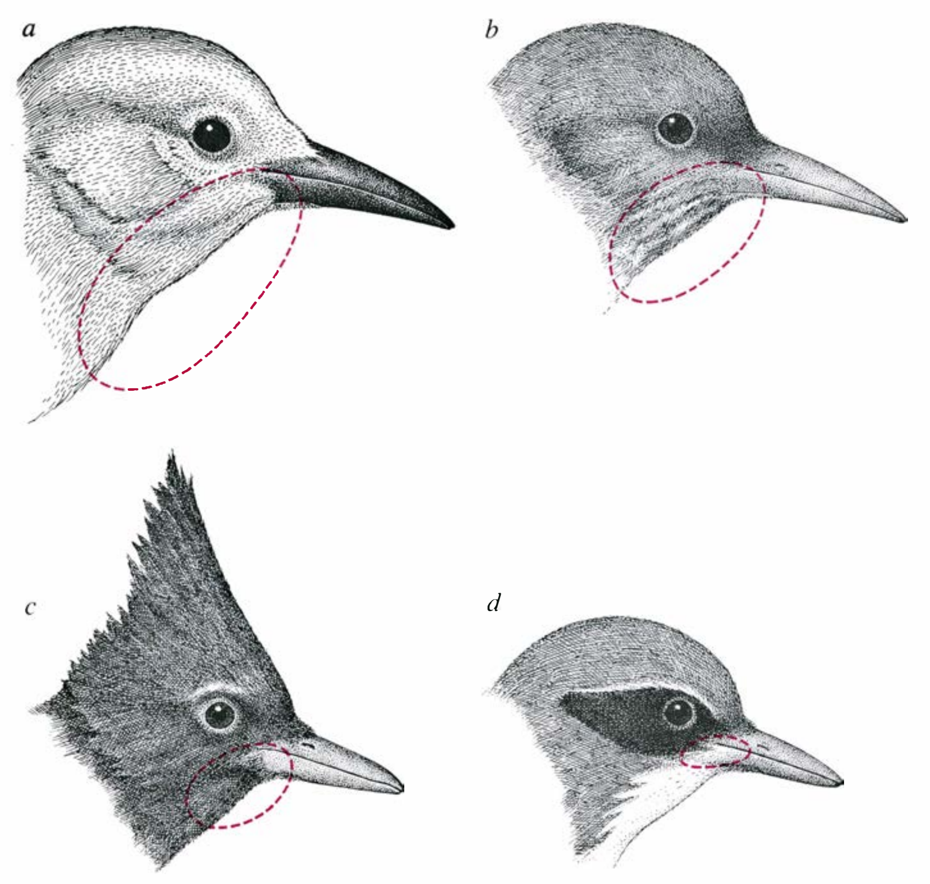


3. BOLSA SUBLINGUAL de un cascanueces de Clark; se trata de una cavidad situada bajo la lengua en la que el ave acarrea las semillas que va a esconder. Se trata, pues, de una especialización anatómica que permite al cascanueces adaptarse bien a su vida de transporte y acaparamiento de piñones. Cuando Vander Wall realizó esta radiografía, la bolsa contenía 38 piñones, cuyo peso neto era de 30,6 gramos.

Una observación más atenta de nuestros datos mostró que un carbonero no iba simplemente a los mismos agujeros en cada prueba. Incluso los agujeros por los que el carbonero mostraba una mayor preferencia tenían más probabilidades de ser inspeccionados durante la recuperación de las semillas si, efectivamente, el párido había depositado alguna en ellos. No obstante, el problema era fundamental. En las pruebas de memoria convencionales realizadas en el laboratorio, el experimentador suministra la información que el animal debe intentar recordar. En nuestros ensayos, las propias aves producían la información al almacenar las semillas donde querían. Cualquier tendencia a escudriñar en los mismos agujeros en visitas sucesivas a la habitación generaría una conducta análoga a la que mostraría si su memoria de almacenamiento fuera buena.

Por ello, nuestro siguiente experimento se diseñó para hacer que la memoria de las aves obrara en contra de su tendencia a preferir determinados lugares de almacenamiento. De nuevo se permitía que el ave almacenara semillas, se la sacaba de la habitación y se la devolvía a ella unas dos horas más tarde. Pero en la segunda sesión el pájaro no tenía que buscar los cañamones, sino almacenar más semillas. Para ello bastaba con dejar un cuenco lleno de cañamones en la habitación. Razonamos que si los páridos recordaban qué agujeros ya tenían semillas, los evitarían al almacenar el segundo lote. Y a la inversa, si inspeccionaban los mismos agujeros cada vez que visitaban la estancia, muchos de los que examinaran en su segunda ronda ya tendrían semilla en su interior.

Los resultados fueron concluyentes. Los pájaros casi nunca inspeccionaban agujeros que contenían semillas al esconder el segundo lote de cañamones. Sin embargo, cuando estaban hambrientos y no había más semillas que almacenar, mostraban su acostumbrada eficacia en recuperar las semillas de los agujeros. Nos interesó señalar que el



4. ESPECIALIZACION ANATOMICA de cuatro aves del sudoeste de los Estados Unidos; se corresponde con su distinto grado de dependencia respecto de la conducta de acaparamiento de alimento. El cascanueces de Clark (*a*) es el ave que más depende de aquél; también es la más especializada: además de una bolsa sublingual tiene un pico largo y aguzado con el que abre las piñas, haciendo palanca, para sacar los piñones. Volador vigoroso, recorre grandes distancias cargado de semillas. El arrendajo piñonero (*b*) y el arrendajo de Steller (*c*) son intermedios, tanto en comportamiento como en anatomía. Transportan semillas en su esófago, que es distendible, y no en una bolsa sublingual especializada al efecto. El arrendajo de matorral (*d*) es, de los cuatro, el que menos depende del almacenamiento de semillas, y es asimismo el de anatomía menos especializada. Puede transportar semillas únicamente en la boca y el pico. En cada dibujo se ha indicado la capacidad de portar piñones mediante una línea discontinua de color.

pájaro que almacenaba un segundo lote de semillas escogía con frecuencia un agujero situado cerca de uno de los que había empleado al almacenar la primera tanda de semillas. Esto es, el ave parecía recordar y evitar agujeros concretos, no sólo determinadas ramas o partes de la habitación.

Tras almacenar dos tandas de semillas, el pájaro podía recuperar todos los cañamones. Podíamos, así, comprobar el efecto de inmediatez. Si los carboneros palustres presentan tal efecto, deberían recuperar las semillas del segundo lote antes, o con más frecuencia, que las del primero. Nuestros resultados indican que existe tal efecto de inmediatez, aunque no es intenso. Mas si la memoria de los lugares de almacenamiento se desvanece lentamente, los dos episodios de acaparamiento tendrían que haberse separado por dos horas largas para que el efecto de inmediatez fuera apreciable.

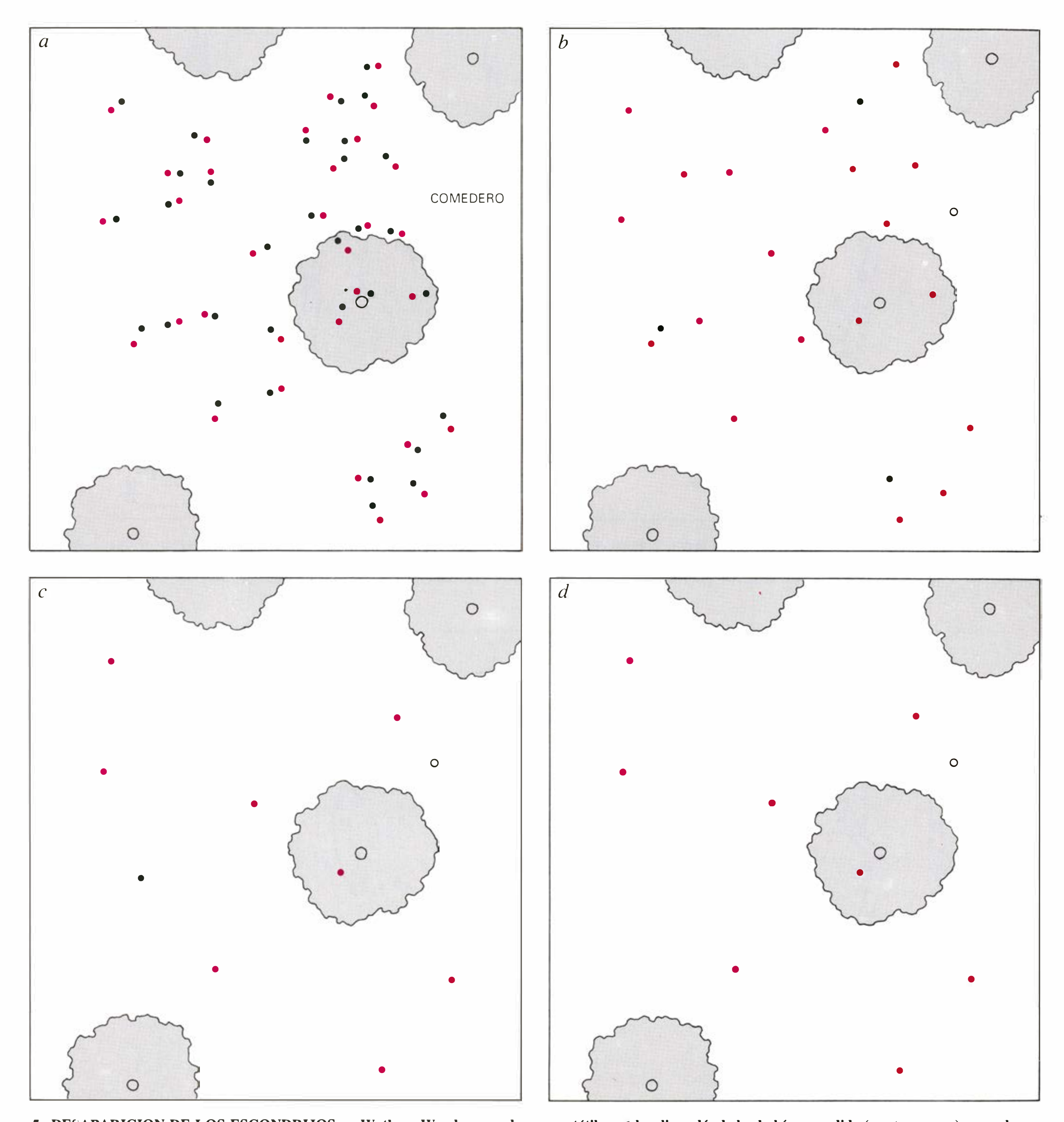
Los experimentos aportaron un nuevo dato. Es evidente que un buscador eficaz de alimento almacenado rara vez ha de inspeccionar lugares vacíos. Más aún, ha de poder recordar dónde ha ido a recuperar alimento para no repetir la visita. Las ratas de laboratorio que se enfrentan a la tarea de recolectar alimento en distintos puntos de un gran laberinto gozan de esta última capacidad. Los carboneros palustres también la poseen. En los experimentos que realizamos Krebs y yo, los agujeros visitados por un pájaro mientras recuperaba los cañamones recibían una segunda inspección con mucha menos frecuencia de lo que se esperaría si el párido no recordara y evitara tales lugares. En experimentos realizados por Sherry, los pájaros tendían a no explorar zonas de musgo que ya habían inspeccionado en busca de semillas. El carbonero palustre en busca de alimento almacenado utiliza, pues, dos tipos de información: recuerda dónde ha almacenado comida y qué lugares ha inspeccionado ya.

Si el carbonero palustre almacena en su cerebro la información sobre los lugares de acaparamiento, su memoria deberá ajustarse a una peculiaridad del sistema nervioso que parece darse en muchas aves. En muchos vertebrados (hombre incluido) la información que percibe la mitad izquierda del campo visual procede de ambos ojos, pero se dirige a la mitad derecha del cerebro, y viceversa. La situación es algo distinta en aves como el carbonero palustre, en las que los ojos se hallan situados a los lados de la cabeza y cada ojo inspecciona un campo visual distinto. El seguimiento de las rutas de información en el cerebro del ave y los resultados de experimentos, realizados principalmente con palomas, sugieren que en estas aves la información que llega al cerebro procedente de un ojo se almacena mayoritariamente en la mitad del cerebro del lado opuesto de la cabeza. Para decirlo de otro modo, estas aves presentan muy poca o ninguna transferencia interocular para muchos tipos de información. Sherry, trabajando con Krebs y Cowie, se valió de este hecho para interpretar un curioso comportamiento que muestran los carboneros palustres cuando almacenan una semilla. Después de haber colocado la semilla en su sitio, mediante unos golpecitos o ligeros empujones, el párido inclina su cabeza primero hacia un lado y luego hacia el otro. Es como si estuviera mirando el lugar de almacenamiento o los hitos situados en torno al mismo, primero con un ojo y luego con el otro.

¿Inclina el carbonero palustre la cabeza para almacenar información visual sobre los lugares de acaparamiento en ambas mitades del cerebro? Sherry se aprestó a comprobarlo: dejó que carboneros palustres almacenaran pipas de girasol en el lugar de su elección, en una gaveta de musgo; a cada pájaro se le había adherido una funda translúcida sobre un ojo. Si posteriormente se permitía al ave buscar las semillas con el mismo ojo cubierto, el animal encontraba las pipas sin mayores dificultades. Si cambiábamos el parche de ojo, de modo que tuviera que buscar con el ojo que no había visto los lugares de almacenamiento, el animal se comportaba como si no recordara dónde estaban las semillas. Si los carboneros palustres carecen de transferencia interocular de información, o la presentan reducida, este comportamiento constituye una prueba más de que utilizan su memoria para recuperar sus provisiones.

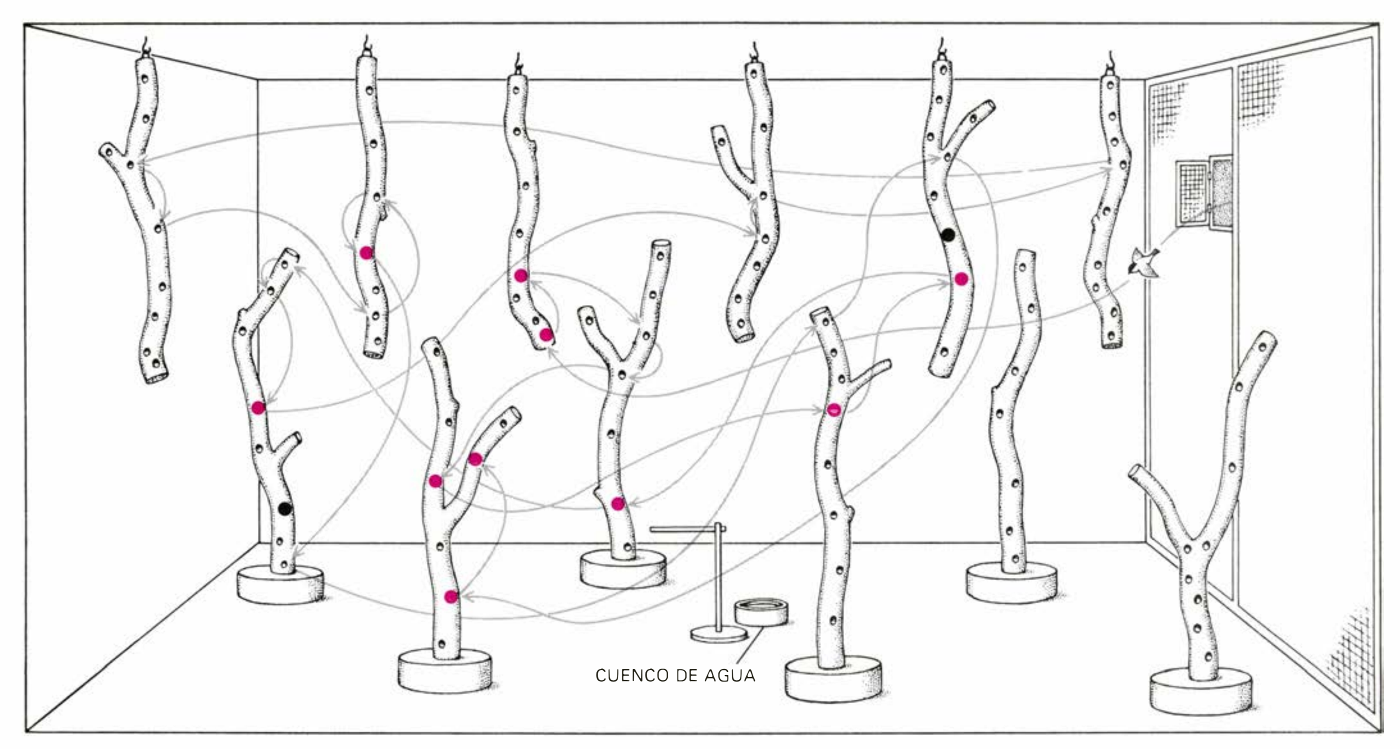
¿Ocurre igual en otras aves que almacenan alimento? En los últimos años, varios investigadores han estudiado aves del tipo del cascanueces de Clark, que establece miles de escondrijos de provisiones invernales. También aquí se han buscado pruebas de memoria en la naturaleza y en el laboratorio. Un estudio de campo que intente descubrir extremos tales como si el cascanueces de Clark recupera sus propias provisiones presenta dificultades obvias: el gran número de escondrijos que se establecen y el largo intervalo que media entre almacenamiento y recuperación. Pocas personas estarán dispuestas a permanecer varios meses junto a un escondrijo de cascanueces con el ánimo de averiguar si el ave que lo enterró torna.

Sin embargo, según Diana Tomback, de la Universidad de Colorado en Denver, podía obtenerse mucha información observando la conducta de los cascanueces de Clark cuando buscaban escondrijos. Tomback sacó partido al

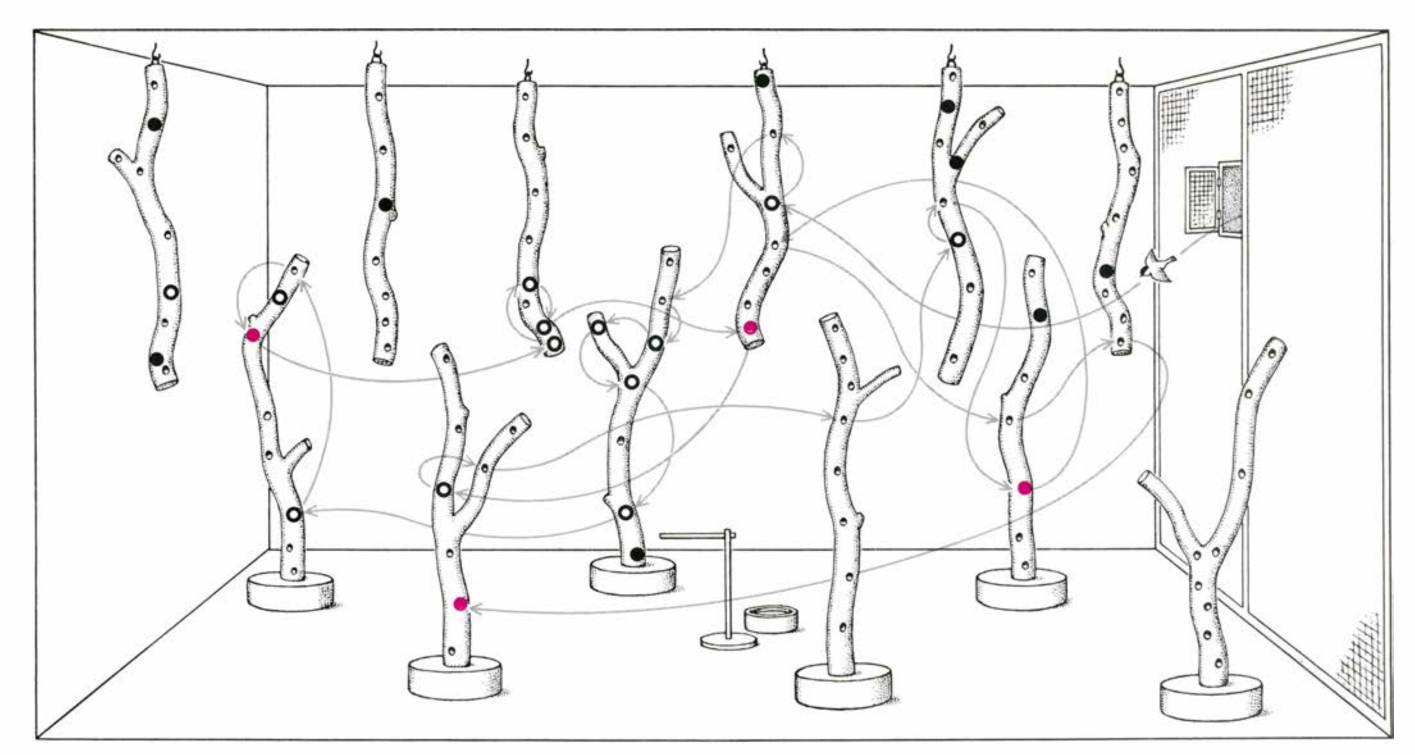


5. DESAPARICION DE LOS ESCONDRIJOS en Wytham Wood, cerca de Oxford; ésta fue la primera prueba de que los carboneros palustres utilizaban la memoria para recuperar alimento acaparado. Richard J. Cowie, John R. Krebs y David F. Sherry, de la Universidad de Oxford, recubrieron 30 pipas de girasol con una sustancia radiactiva y emplearon un contador de centelleo portátil para localizar dónde las había escondido (*puntos negros*) un carbonero palustre (*a*). A un metro de distancia de cada escondrijo establecieron una falsa reserva (*puntos de color*). Diez horas más tarde (*b*), 19 escondrijos falsos permanecían intactos, por sólo tres de los del pájaro. Los investigadores hicieron comprobaciones posteriores a las 19 horas (*c*) y a las 30 horas (*d*).

dato siguiente: las aves dejaban un registro de sus pesquisas, en forma de picotazos, en la nieve y en la tierra donde excavaban. Además, las pesquisas que tienen éxito pueden identificarse por la presencia de cáscaras de piñón cerca de los agujeros de los que se tomaron las semillas. Si los cascanueces de Clark buscan al azar, los agujeros en los que los sondeos han tenido éxito y los grupos de agujeros infructuosos debieran hallarse distribuidos más o menos uniformemente sobre el territorio sometido a inspección por las aves. Por el contrario, Tomback encontró que los tanteos infructuosos se encontraban arracimados alrededor de los fructíferos. La pauta de distribución indicaba que los cascanueces no buscaban los escondrijos por tanteo. Además, a principios de primavera (antes de que los roedo-



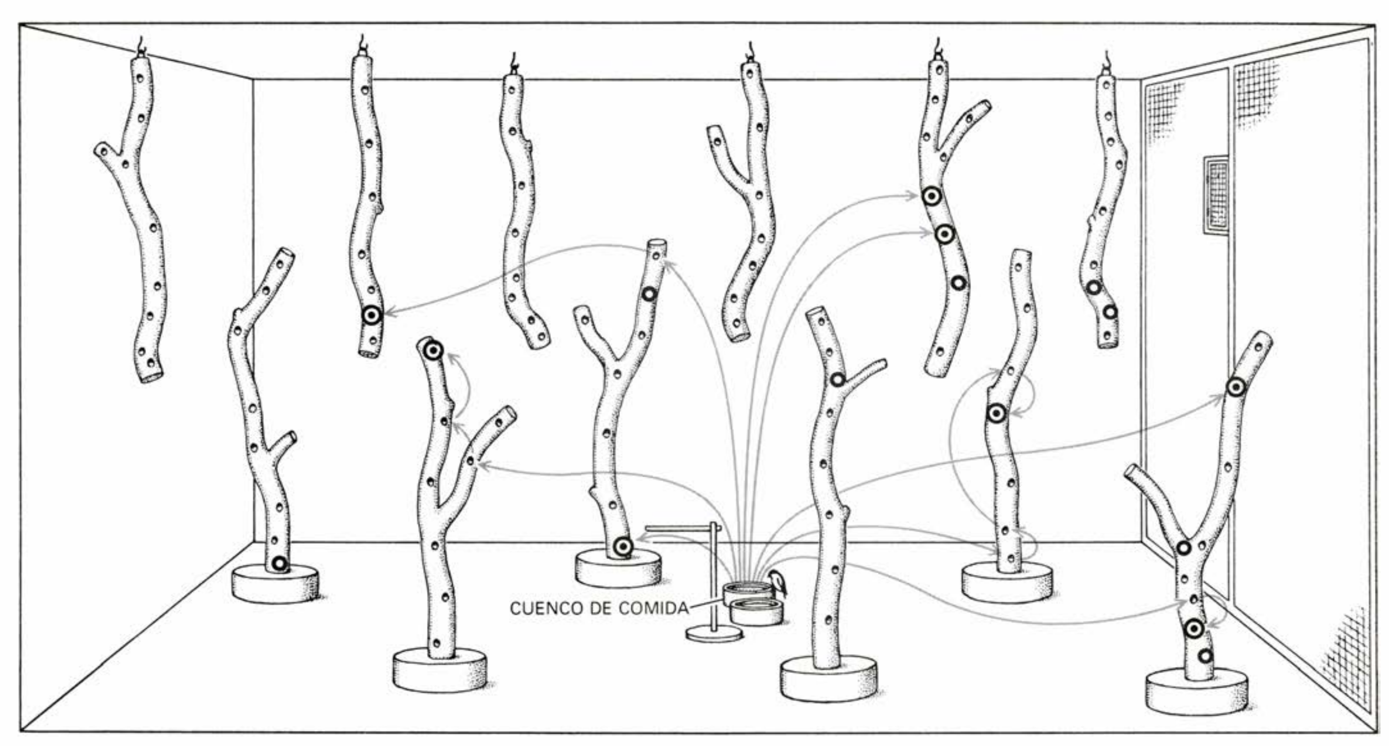
6. CUATRO EXPERIMENTOS para comprobar la memoria de los carboneros palustres, realizados en colaboración con Krebs. En cada caso, un pájaro almacenaba semillas en algunos de los agujeros que los experimentadores habían perforado en ramas de árbol. Aquí, un carbonero palustre acaparó 12 semillas (*puntos negros*). Readmitido a la habitación dos horas y media más tarde, encontró 10 (*puntos de color*) inspeccionando tan sólo 24 agujeros.



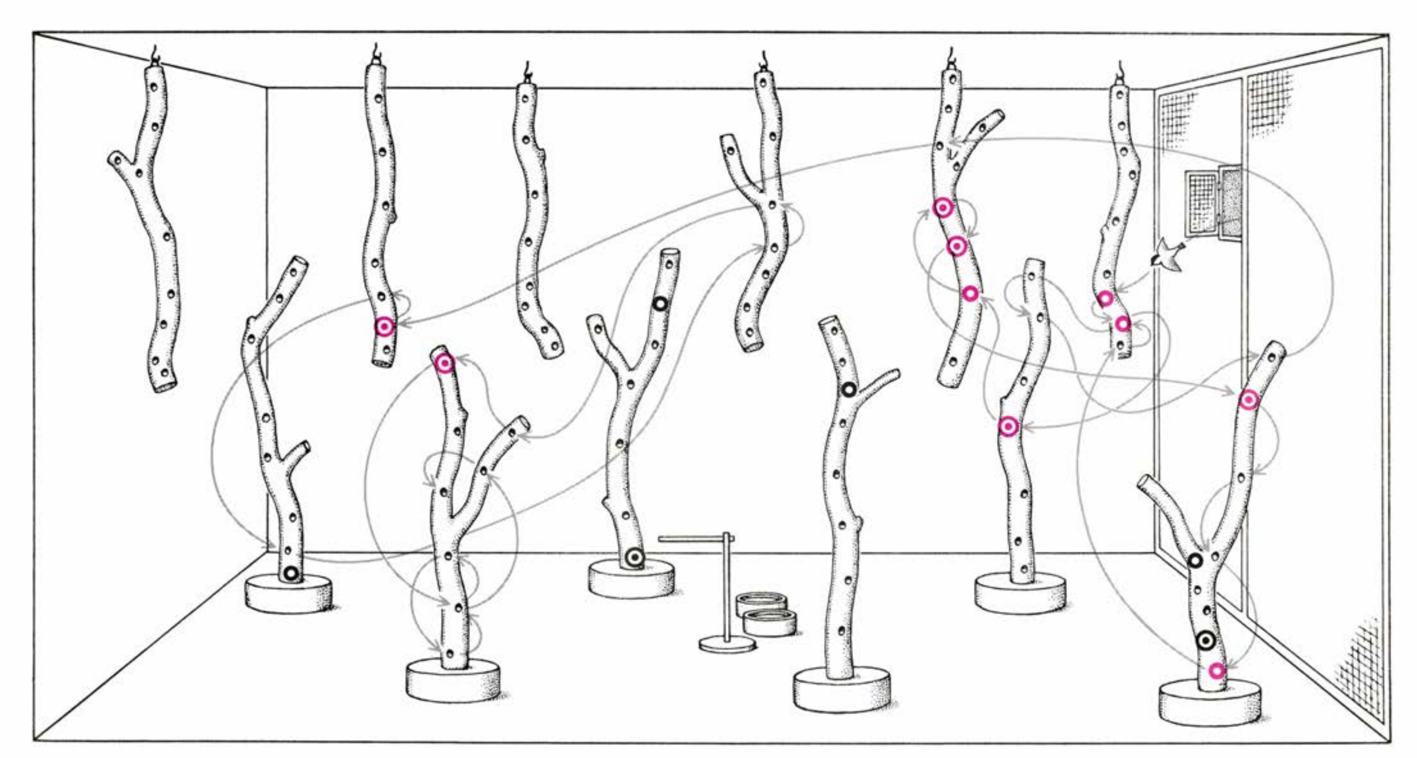
7. EN UN SEGUNDO EXPERIMENTO un carbonero palustre almacenó 13 semillas (*círculos negros*), pero Krebs y el autor las desplazaron luego a otros agujeros (*puntos negros*). De los primeros 24 agujeros que el ave inspeccionó cuando fue devuelta a la habitación, 11 eran agujeros que había empleado antes. El experimento indica que el ave no estaba detectando las semillas mediante alguna señal, como su olor. El pájaro logró encontrar cuatro semillas.

res localizaran muchos de los escondrijos), alrededor de los dos tercios de las pesquisas habían tenido éxito, mucho más de lo que cabría esperar de una búsqueda al azar.

Las observaciones de Tomback no prueban de manera inequívoca que las aves recordaran dónde se hallaban sus escondrijos. Podían haber olido sus provisiones (aunque, de nuevo, es ésta una posibilidad poco verosímil). Alternativamente, podían haber buscado principalmente en los lugares más propicios para albergar escondrijos. Recientes estudios de laboratorio, llevados a cabo por Stephen B. Vander Wall en la Universidad estatal de Utah, parecen descartar estas posibilidades. Vander Wall estudió cuatro cascanueces de Clark cautivos en un gran aviario dispuesto al aire libre. Dos de las aves



8. DOS LOTES DE SEMILLAS pudo acaparar el carbonero palustre en el tercer experimento. Al escoger agujeros para el segundo lote (*círculos negros con puntos*) evitó los agujeros en los que había almacenado el primer lote (*círculos negros*). El experimento estableció que el pájaro no se limitaba a visitar un determinado grupo de agujeros, ya estuviera almacenando semillas o recuperándolas. El ave podía recordar por dónde había pasado antes.



9. RECUPERACION DE AMBOS LOTES de semillas por parte del carbonero palustre que las había acaparado; éste fue el experimento final de la serie. Inspeccionando 29 agujeros encontró 10 semillas: cuatro (*círculos de color*) del primer lote de ocho y seis (*círculos de color con puntos*) del segundo. Sus siete primeras inspecciones de agujeros se vieron todas coronadas por el éxito. Sólo en una ocasión cometió el error de volver a visitar un agujero.

escondieron inmediatamente semillas en el suelo arenoso del cercado; las otras dos, no. Las aves no acaparadoras excavaban en la arena con su pico y se comían las semillas que lograban descubrir. Vander Wall dejó que los dos acaparadores enterraran semillas en el aviario; permitó luego que las cuatro aves buscaran las semillas.

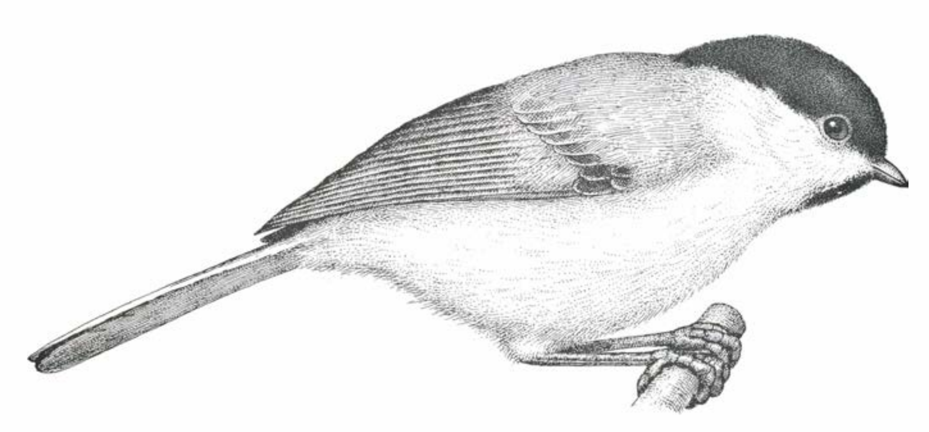
Los resultados demuestran espectacularmente el empleo de la memoria. Cuando los cascanueces que habían escondido semillas excavaron en la arena para encontrarlas, el 70 por ciento de sus pesquisas tuvieron éxito. Cuando lo hicieron los no acapadores, encontraron semillas sólo en el 10 por ciento de los intentos. De hecho, un índice de éxitos del 10 por ciento es bastante superior al que se esperaría de un ave que buscara de forma aleatoria. Por otra parte, los dos cascanueces que escondían semillas preferían situar sus escondrijos cerca de objetos prominentes del aviario, como troncos y piedras, y las cuatro aves buscaban con más frecuencia en las inmediaciones de estos objetos. Esta tendencia explicaría el éxito de los no acaparadores.

Quizá los objetos prominentes de las inmediaciones sirvan de señales a la memoria cuando las aves van en busca de sus almacenes. Para probar esta hipótesis, Vander Wall cubrió el suelo del aviario con una lámina de plástico, dejando descubierta únicamente una zona ovalada de la que se había cortado la lámina. A cada extremo del óvalo dispuso cuatro objetos grandes a modo de rocas. Los cascanueces se procuraron sus escondrijos donde quisieron, dentro del óvalo descubierto. Se prolongó luego 20 centímetros un extremo del óvalo y se desplazaron los cuatro objetos de aquel extremo en la dirección del alargamiento, de modo que se hallaran también a 20 centímetros de su posición original.

¿Lograrían los cascanueces encontrar ahora los escondrijos en ambos extremos del óvalo, o quizá sus intentos quedarían desplazados 20 centímetros en el extremo prolongado? Ocurrió esto último. Hacia el extremo inalterado del óvalo, el índice de éxitos de las aves fue alto; hacia el extremo alterado, el índice de éxitos bajó. La mayoría de intentos en el extremo alterado se hallaban a unos pocos centímetros del punto teóricamente "correcto" respecto de la posición alterada del objeto grande más cercano. Las aves también cometieron errores hacia la parte media del óvalo, pero de menor calibre. Este resultado sugiere que los lugares de almacenamiento se habían establecido con referencia a ambos grupos de objetos.

Cabe preguntarse si una especie que dispersa sus provisiones y fía en la memoria para recuperarlas hace algún otro uso de esa facultad. Es decir, ¿están mejor dotadas las especies que almacenan alimento para recordar la localización de otros objetos? Para hallar la respuesta habría que someter a los animales a pruebas de memoria espacial que no tuvieran que ver con el alimento almacenado. No se han efectuado aún tales ensayos con aves que almacenan alimento, pero la propia reflexión sobre su diseño centra nuestra atención en varios puntos de especial interés.

Entre las técnicas de laboratorio empleadas por los psicólogos que investigan la memoria de los animales, el análogo más cercano al acaparamiento de alimento es un experimento de respuesta diferida. En este tipo de experimento, a un animal se le muestra alguna señal visual, como una luz que emite destellos o una figura geométrica. Posteriormente se le ofrece elegir entre la señal original y otra. Se le premia si opta por la que vio primero. La actuación típica de una rata de laboratorio o de una paloma cae al nivel de azar si transcurren unos cuantos segundos o minutos entre la primera prueba visual y la segunda.

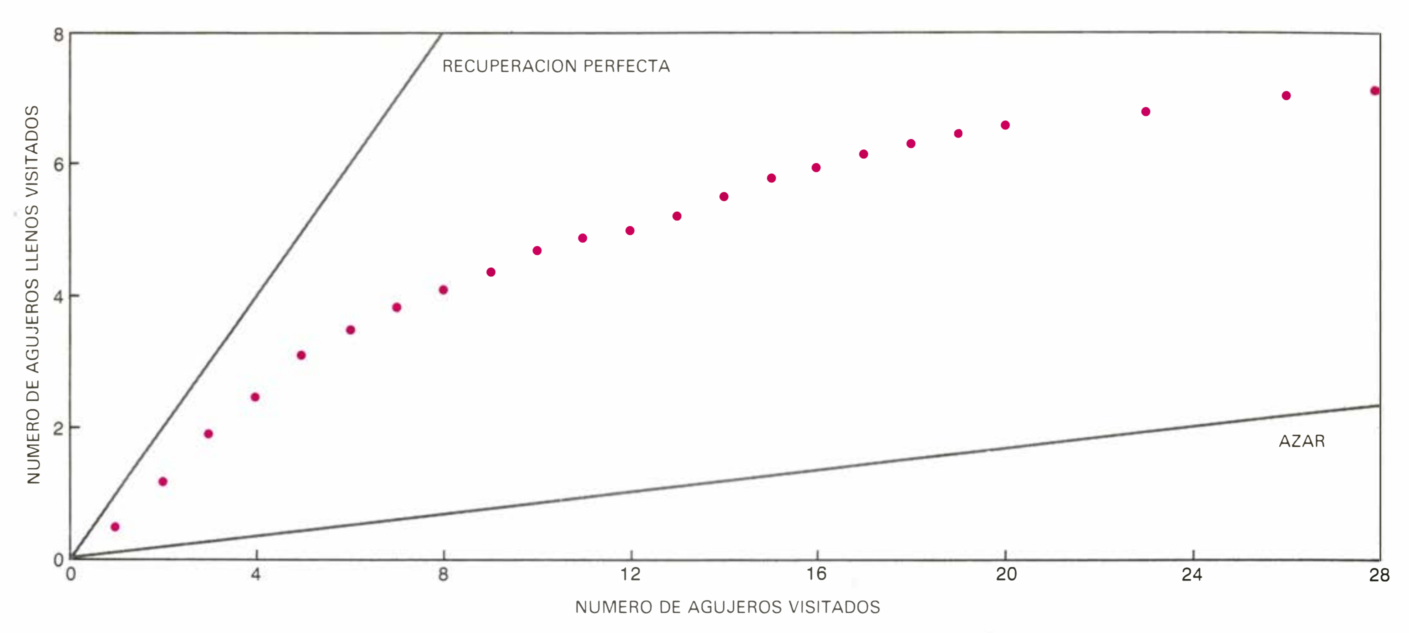


10. CARBONERO PALUSTRE; es un pequeño párido europeo que esconde semillas si dispone de más de las que precisa de inmediato para su alimentación. El carbonero palustre retorna a recuperar sus acúmulos al cabo de unas pocas horas o, todo lo más, unos pocos días después de haberlos depositado.

Recientemente, Donald M. Wilkie y Russel Summers, de la Universidad de Columbia Británica, adiestraron palomas en una variante del procedimiento de respuesta diferida que exige memoria espacial. La investigación ilustra algunas de las diferencias entre los estudios de laboratorio convencionales sobre memorial animal y aquellos otros en los que se trabaja con acaparadores de alimento. En el experimento de Wilkie y Summers, una paloma se enfrentaba a un conjunto de nueve discos blancos, dispuestos en un cuadro de tres por tres, que podían iluminarse individualmente por detrás. Al inicio de cada prueba se iluminaba brevemente un disco. Unos pocos segundos más tarde se iluminaba el disco de nuevo, junto con uno o varios discos más. Se premiaba a la paloma (recibía una bolita de comida) si picoteaba el disco que se había iluminado dos veces.

La tarea de la paloma parece ser muy similar a la que se le presenta a un ave que almacena comida. Después de todo, un ave que almacena alimento se halla expuesta a señales visuales en el lugar de almacenamiento, y para recuperar el alimento debe seleccionar el lugar entre otros similares de la vecindad. Y, sin embargo, no teniendo que recordar más que una señal cada vez, las palomas del experimento de Wilkie y Summers parecían no saber qué disco se había iluminado a los pocos segundos después de haberlo visto; por contra, el cascanueces de Clark recuerda miles de lugares de almacenamiento durante meses.

La conducta de la paloma difería en varios aspectos de la del cascanueces. La paloma observaba pasivamente la señal que después debía recordar. Las señales se hallaban invariablemente a unos pocos centímetros de distancia unas de otras y diferían sólo en cuanto a su posición en el dispositivo experimental; la comida con la que se premiaba a la paloma no se hallaba detrás de la señal propiamente dicha. Cabe suponer que cualquiera de estos factores hiciera la tarea de la paloma más difícil que la labor a la que se enfrenta un ave que acapara alimento. Por ejemplo, investigaciones de laboratorio realizadas principalmente con ratas demuestran



11. ACTUACION DEL CARBONERO PALUSTRE, según los resultados de los experimentos efectuados por Krebs y el autor. Si el pájaro hubiera buscado al azar ocho semillas en 100 agujeros, habría inspeccionado 12 agujeros para encontrar cada semilla. En cambio, con una memoria sólida inspeccionaría sólo un agujero por semilla. Tuvo una actuación intermedia: en 12 pruebas inspeccionó una media de 3,7 agujeros para descubrir cada semilla.

que los animales pueden recordar, por varias horas, los lugares que visitan.

Por tanto, quizá sean los tipos de señales disponibles en los lugares de acaparamiento lo que los haga memorables. Según esta hipótesis, un ave acaparadora de alimento no tendría por qué poseer necesariamente una memoria superior, en mucho, que la de otros animales. Al adoptar una vida de acaparamiento de alimento, se habría puesto en una situación en la que su memoria resulta de particular utilidad. ¿Cómo comprobarlo? Las comparaciones de memoria entre especies que acaparan alimento y especies que no lo hacen se hallan con frecuencia comprometidas por diferencias en las motivaciones y en las capacidades sensoriales y motrices de las especies. Resulta obvio que las comparaciones serían más significativas si las especies objeto de estudio guardaran un estrecho parentesco.

A este respecto conviene resaltar que, entre los Córvidos (arrendajos, cuervos, cornejas y cascanueces) y también entre los Páridos, algunas especies almacenan alimento mientras que otras no lo hacen. Entre los Páridos, por ejemplo, los herrerillos capuchinos y los carboneros sibilinos parecen almacenar comida en otoño y recurren a ella a lo largo del invierno, una pauta que recuerda el comportamiento de los cascanueces de Clark. Los carboneros palustres almacenan comida en exceso para intervalos de tiempo relativamente cortos. Los herrerillos comunes y los carboneros comunes no guardan nada.

Vander Wall, trabajando con Russell P. Balda, de la Universidad de Arizona septentrional, ha descrito, en los Córvidos, un ejemplo de varias especies emparentadas que difieren en su tendencia a acaparar. El cascanueces de Clark vive en las mismas zonas de los Estados Unidos occidentales que otros tres miembros de la familia Córvidos, a saber, el arrendajo piñonero, el arrendajo de Steller y el arrendajo de matorral. Las cuatro especies difieren mucho en su dependencia de piñones almacenados y, concomitantemente, difieren en su habilidad a la hora de recolectar y transportar las semillas.

El cascanueces de Clark es el que muestra mayor dependencia. Gracias a la explotación de los piñones almacenados puede empezar a reproducirse muy temprano, en febrero, mucho antes que la mayoría de las demás aves de la zona. El cascanueces de Clark es asimismo el más especializado, desde el punto de vista anatómico, para la recolección y el acaparamiento de semillas. Además de la bolsa sublingual, en la que transporta gran número de piñones, posee un pico largo y puntiagudo con el que fuerza, a modo de palanca, las piñas verdes antes de que se abran por sí mismas. Además, el cascanueces de Clark es un buen volador. Puede transportar pesadas cargas a grandes distancias, hasta los lugares de almacenamiento adecuados.

El arrendajo de matorral, en cambio, es la menos especializada de las cuatro especies. Carece de las adaptaciones anatómicas del cascanueces de Clark y se basa mucho menos en semillas almacenadas. Los piñones que almacena los esconde enteramente en la zona donde vive, es decir, la zona que defiende como su propio territorio. El arrendajo piñonero y el de Steller quedan en medio, tanto en la anatomía como en el comportamiento. ¿Está menos desarrollada la memoria espacial en estas tres especies que en el cascanueces de Clark? La respuesta requiere ulteriores investigaciones, en el laboratorio y en la naturaleza.

Las investigaciones sobre la memoria y el acaparamiento de alimento que he descrito se sitúan en la interfase entre los estudios zoológicos y psicológicos del comportamiento animal. Aumenta el número de estos ensayos a medida que los psicólogos se convencen del interés de los estudios de laboratorio sobre aprendizaje y memoria para analizar de qué modo resuelven los animales los problemas con los que se enfrentan en la naturaleza. Al mismo tiempo, los zoólogos se están interesando por aspectos del comportamiento de búsqueda del alimento que requieren claramente capacidad de aprendizaje y de memoria. Por ello, los esfuerzos conjuntos pueden combinar la comprensión del zoólogo sobre cómo se adapta el comportamiento al ambiente, con los instrumentos conceptuales y las técnicas experimentales del psicólogo. Tales esfuerzos son el inicio de nuevos y sugestivos logros en el estudio de la memoria.

Reacciones químicas oscilantes

Hubo un tiempo en que se creyó que violaban las leyes de la naturaleza. Hoy, esas reacciones, en que las concentraciones suben y bajan periódicamente, están renovando la química y la biología

Irving R. Epstein, Kenneth Kustin, Patrick De Kepper y Miklós Orbán

Los fenómenos periódicos u oscilantes se presentan en física, astronomía y biología. Van desde el movimiento de los péndulos hasta las órbitas de los planetas y los complejos relojes biológicos que gobiernan el comportamiento diario y estacional de los organismos. Hasta ayer mismo, los químicos creían que las reacciones que se desarrollaban en sus tubos de ensayo y vasos de precipitados eran inmunes a ese tipo de conducta periódica, tan común en otros campos de la ciencia. La inmensa mayoría de los químicos que se formaron antes de 1950 habrían afirmado que cualquier mezcla de sustancias inorgánicas sencillas que participara en una reacción que oscilara visible y periódicamente violaría una ley inmutable de la naturaleza. Incluso hoy damos por sentado que las reacciones químicas discurren en una dirección. Si dos sustancias reaccionan para producir una tercera, la reacción continuará, se supone, hasta que se consuman los reactivos o hasta que se alcance el estado de equilibrio. Nadie espera, por lo común, que la concentración de los productos intermedios llegue a cierto nivel, caiga luego a otro más bajo, torne a subir para bajar de nuevo, una y otra vez, hasta que resulten productos estables, resistentes a más cambios.

Con carácter esporádico se mencionaba esta clase de reacciones en la literatura química del siglo XIX y comienzos del XX; no merecían, sin embargo, la ponderación de los profesionales: se las suponía fenómenos no reproducibles, atribuibles quizás a procesos extraños como la corrosión o la formación de películas durante el curso de la reacción. Pero a lo largo de los últimos 25 años el estudio de las reacciones químicas oscilantes se ha rodeado de respetabilidad suficiente para constituir hoy una de las ramas más inquietas de la química. De su avance se esperan nuevos planteamientos de la dinámica química y el mecanismo de la catálisis, y, posiblemente también, nuevas claves para entender fenómenos periódicos que se observan en biología y geología y carecen todavía de explicación. Las primeras reacciones químicas oscilantes se descubrieron por azar. Ahora sabemos ya las condiciones que conducen a la oscilación y nos permiten montar una búsqueda sistemática de nuevos sistemas oscilantes. Una búsqueda que se ha ganado ya la recompensa.

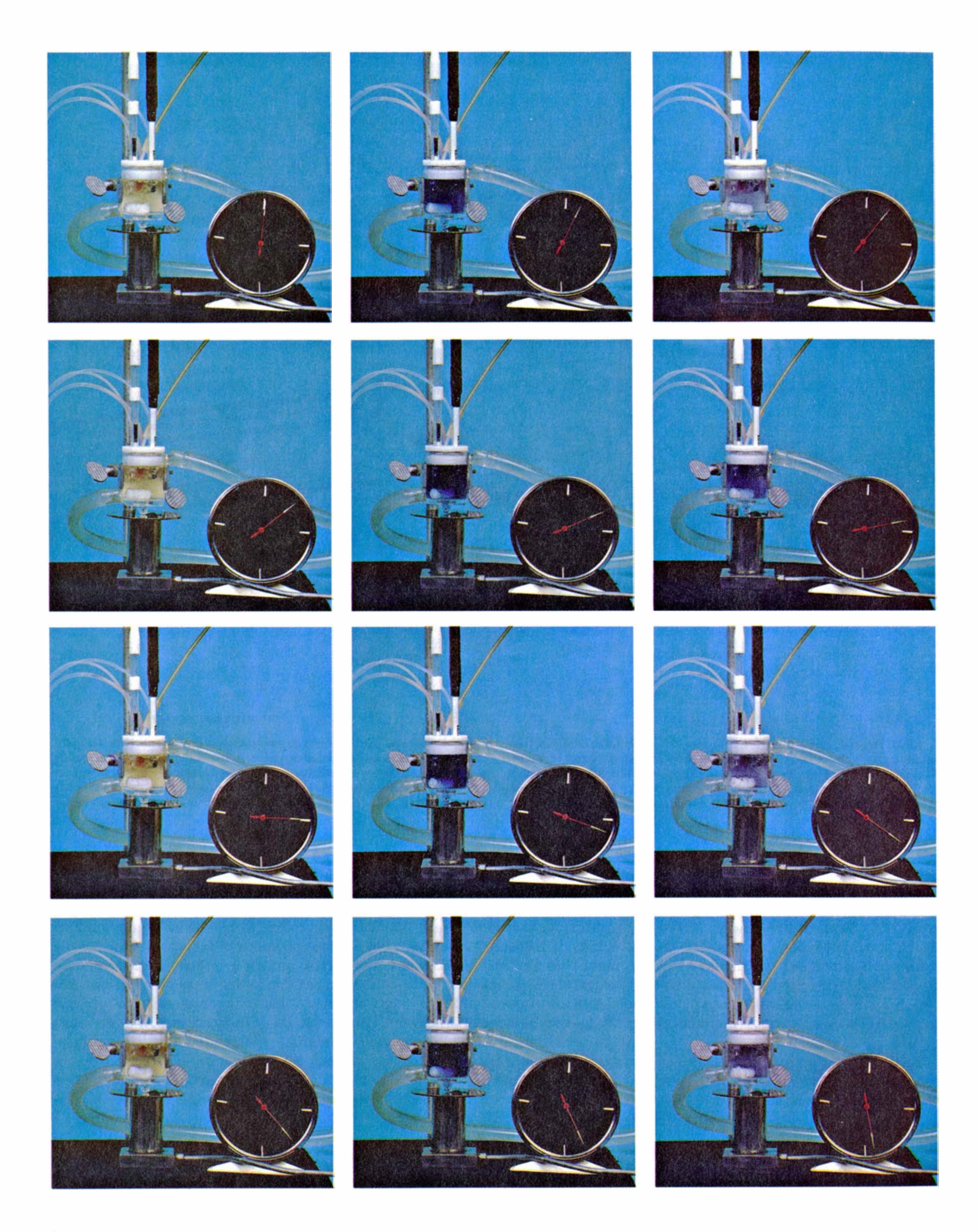
Los químicos reacios a aceptar la realidad de las reacciones oscilantes se fundaban, principalmente, en la segunda ley de la termodinámica. En su formulación mejor conocida, establecida en el siglo pasado por el físico alemán Rudolf Clausius, determina la segunda ley que la entropía o el desorden del universo tiende a aumentar. Aplicado a las reacciones químicas, este principio exige que un sistema químico debe aproximarse continuamente hacia un estado final de equilibrio, siempre que no se introduzca materia o energía. Esto es, si A se transforma en B, ha de hacerlo siguiendo el curso de la reacción y no volver de nuevo a A durante el proceso. Si alguna reacción parecía violar esta regla, había que buscar la razón en las condiciones experimentales, pobremente controladas, cuando no trampas deliberadas, puesto que vendría a constituir una suerte de máquina química en movimiento perpetuo. Esa era la opinión reinante.

Cabe suponer, pues, con cuánta indiferencia se recibió el trabajo sobre una reacción oscilante que publicó en 1921 William C. Bray, de la Universidad de California en Berkeley. Investigaba Bray la conducta del iodato, el ion del iodo oxidado (IO_3^-); al catalizar la descomposición del peróxido de hidrógeno (agua oxigenada) en agua y oxígeno, observó que, bajo ciertas condiciones, la velocidad de producción de oxígeno y la concentración de iodo en la disolución cambiaban periódicamente. A lo largo de los cincuenta años subsiguientes, nadie reparó en esa observación. Los contados trabajos que se ocuparon de la reacción de Bray explicaban las oscilaciones como artificios creados por partículas de polvo o impurezas; pero faltó quien examinara el fenómeno con mayor profundidad.

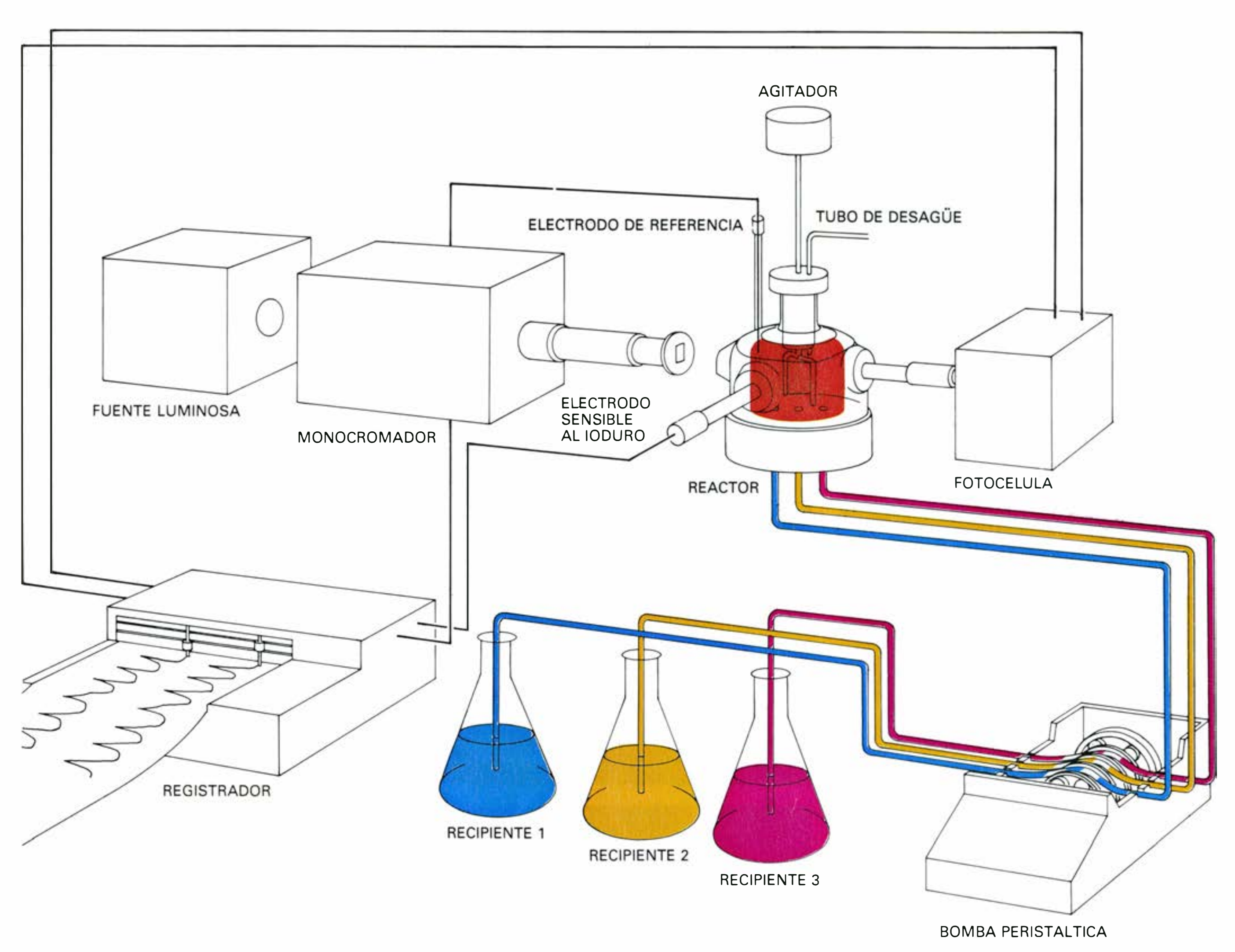
Lo veremos después, el iodo, el ion ioduro (I^-) y el ion iodato desempeñan un papel destacado en varios sistemas oscilantes de reciente creación. La época moderna de las reacciones químicas oscilantes comienza con un descubrimiento, obtenido al azar en 1958, por el químico ruso B. P. Belousov. Observó que si disolvía, en agua, ácido cítrico y ácido sulfúrico con bromato potásico y una sal de cerio, el color de la mezcla cambiaba periódicamente de incoloro a amarillo pálido. En 1958 había ya quimicofísicos dispuestos a considerar posible esta información; por dos razones: primero, porque la reacción de Belousov se podía repetir con facilidad y, segundo, porque las ideas termodinámicas del siglo XIX habían gozado de un amplio desarrollo en los años que siguieron a la segunda guerra mundial.

Adelantado de este empeño fue Ilya Prigogine, de la Universidad Libre de Bruselas, quien advirtió que la termodinámica clásica de Clausius no sólo requería que los sistemas se hallaran aislados de su entorno, sino además que estuvieran cerca de su estado de equilibrio. Para sistemas lejos del equilibrio, porque la reacción comience o porque el sistema sea abierto, y sometido por tanto a un flujo externo de energía o materia, Prigogine y sus colaboradores desarrollaron una termodinámica irreversible. Por su trabajo en esta línea, Prigogine recibió el premio Nobel de química en 1977.

En los sistemas que están lejos del equilibrio, surgen algunos fenómenos nuevos que se engloban en las llamadas estructuras disipativas. Estas estructuras presentan oscilaciones periódicas en las concentraciones de especies intermedias de una reacción química; pero ni los reactivos iniciales ni los produc-



1. EXHIBICION DE UN SISTEMA QUIMICO OSCILANTE, en el laboratorio de los autores en la Universidad norteamericana de Brandeis. La reacción se desarrolla en una vasija de vidrio, en la que se bombean tres soluciones a velocidad constante; a saber: iodato potásico, una disolución de ácido perclórico y peróxido de hidrógeno y una disolución de ácico malónico que contiene sulfato manganoso monohidratado. El almidón sirve de indicador; forma un complejo azul oscuro en presencia de iodo. De una fotografía a la siguiente, comenzando por la parte superior izquierda y leyendo de izquierda a derecha, media un intervalo inferior a tres segundos (como indica la manilla del segundero del reloj). Esta reacción la idearon dos profesores de bachillerato de California, Thomas S. Briggs y Warren C. Rauscher. La reacción se ha modificado para que se desarrolle en condiciones de flujo continuo.



2. REACTOR DE FLUJO CONTINUO en un tanque agitado; en él buscaron los autores reacciones químicas oscilantes. A tal fin, establecieron las condiciones de no equilibrio necesarias para que se desarrollaran las reacciones en cuestión. La ilustración muestra un experimento típico. Los reactivos, que se albergan en tres depósitos, se bombean continuamente hacia la base de la vasija de reacción. En el primer sistema oscilante que descubrieron los autores, los materiales con que lo alimentaban eran soluciones acuosas ácidas de un clorito, un iodato y un arsenito. Los cambios de color periódicos se seguían con una fotocélula. Las oscilaciones registradas en la concentración de iones específicos (en este caso el iodo) se detectan mediante un electrodo.

tos finales se hallan sujetos a oscilaciones. Los organismos suministran los ejemplos más interesantes y variados de osciladores abiertos y alejados del equilibrio. Se mantienen en un estado de no equilibrio con la admisión de reactivos (alimentos) del medio exterior y la salida de los productos de desecho. Si se detiene cualquier flujo de éstos, cesan las oscilaciones y muere el organismo.

En 1958, las implicaciones de la termodinámica irreversible de Prigogine las entendían sólo un puñado de químicos. Además, del descubrimiento de Belousov no se enteró casi nadie ya que apareció publicado en una oscura colección rusa de resúmenes técnicos sobre medicina nuclear. La reacción oscilante de Belousov presentaba ventajas importantes sobre el anterior sistema de Bray. Se desarrollaba a temperatura ambiente y no generaba productos nocivos. Además, las oscilaciones se veían a simple vista, conforme el cerio cambiaba, en un sentido y otro, de la forma oxidada (cérica), que es amarilla, a la menos oxidada (cerosa), incolora.

Algunos años después de su publicación, la reacción de Belousov llegó a manos de A. M. Zhabotinsky, del Instituto de Biofísica, con sede en los aledaños de Moscú. Modificó algo el protocolo de Belousov (sustituyó el cerio por el hierro, que produjo un color azul rojizo más fácil de distinguir) y emprendió el estudio sistemático de la reacción, que pasó a denominarse de BZ (en honor de Belousov y Zhabotinsky). En los primeros años del decenio de los sesenta, Zhabotinsky publicó una nutrida relación de resultados experimentales; entre ellos, el hallazgo siguiente: si dejaba sin perturbar una fina capa de una solución inicialmente uniforme y roja de BZ, aparecían pronto puntitos azules que crecían formando unas estructuras sorprendentes en formas de anillos.

La reacción BZ y la termodinámica irreversible de Prigogine conocieron un rápido auge en las postrimerías de los sesenta y primeros años de los setenta. En todas partes había algún químico fascinado por el fenómeno BZ, pero la gran mayoría seguía considerándolo mera curiosidad, más útil para deslumbrar a los estudiantes en las clases que tema de investigación serio. En poco tiempo, aparecieron muchas variaciones de la reacción BZ que sustituían uno o varios componentes con especies químicas relacionadas; surgían así osciladores más lentos, más rápidos o de distinto color. La variante más espectacular, desde el punto de vista cromático, es un híbrido de los sistemas BZ y Bray que inventaron dos profesores de bachillerato de San Francisco en 1973, Thomas S. Briggs y Warren C. Rauscher. Su sistema, que utiliza peróxido de hidrógeno, iodato potásico, ácido perclórico, ácido malónico, sulfato de manganeso y almidón, cambia de inco-

loro a dorado y pasa a azul, para volver a empezar. (Para mayores detalles de su trabajo, véase la sección de "Taller y laboratorio" de *Investigación y Ciencia*, septiembre de 1978.)

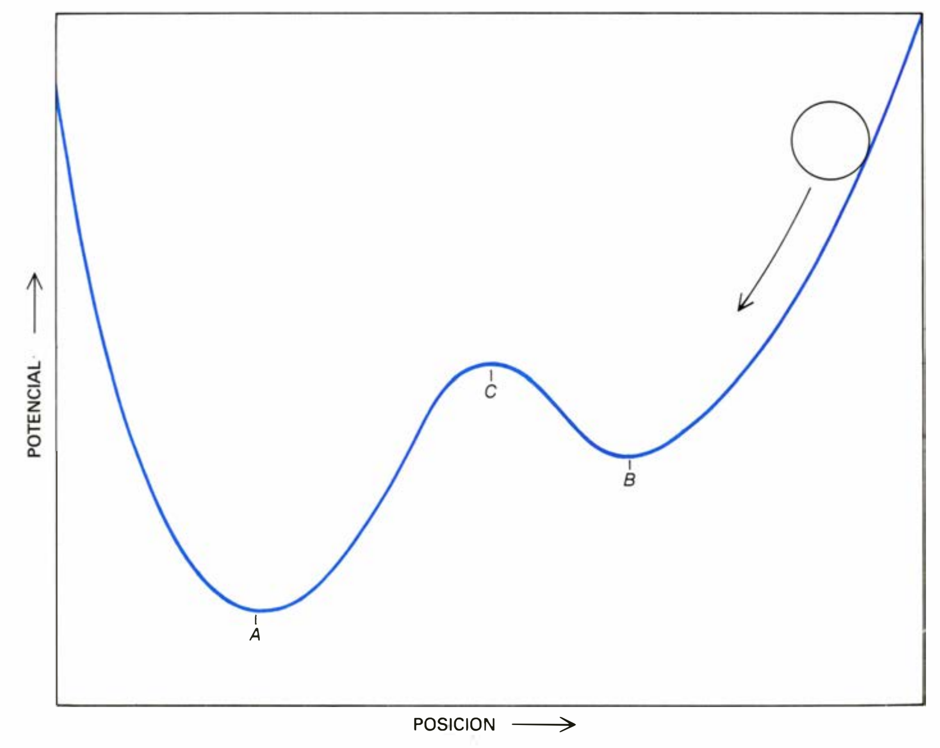
Aunque la reacción BZ y sus variantes gozaban ya de amplia difusión en los últimos años del decenio de los sesenta a través de lecciones y conferencias, la naturaleza del fenómeno oscilante permanecía envuelta en el misterio. Los químicos entienden las reacciones cuando son capaces de exponer el mecanismo subyacente. Un mecanismo es una serie de reacciones componentes, los así llamados pasos elementales, cada uno de los cuales describe un encuentro o colisión real entre moléculas; a través de esas etapas los reactivos se transforman en productos. Las ecuaciones estequiométricas, de las que suelen echar mano los químicos para simbolizar una reacción, ofrecen sólo el resultado neto; pero no dicen nada de cómo se desarrolla la reacción a escala molecular. Por ejemplo, cuando el acetileno (C_2H_2) se combina con oxígeno (O_2) en el soplete de un soldador para formar dióxido de carbono (CO_2) y agua (H_2O), la reacción total se puede escribir $2\ C_2H_2 + 5\ O_2 \rightarrow 4\ CO_2 + 2\ H_2O$.

Es improbable que dos moléculas de acetileno y cinco de oxígeno choquen a la vez y se separen dando cuatro moléculas de dióxido de carbono más dos de agua. En la realidad, la reacción se desenvuelve a través de una secuencia de reacciones elementales; cada una de éstas implica una colisión entre dos moléculas o la ruptura de una de ellas, creando o destruyendo especies intermedias que no aparecen en la ecuación estequiométrica. Descifrar un mecanismo constituye una tarea ardua, para la que se exige ingenio, análisis de muchos datos experimentales y, en no pocos casos, simulación del mismo en el ordenador.

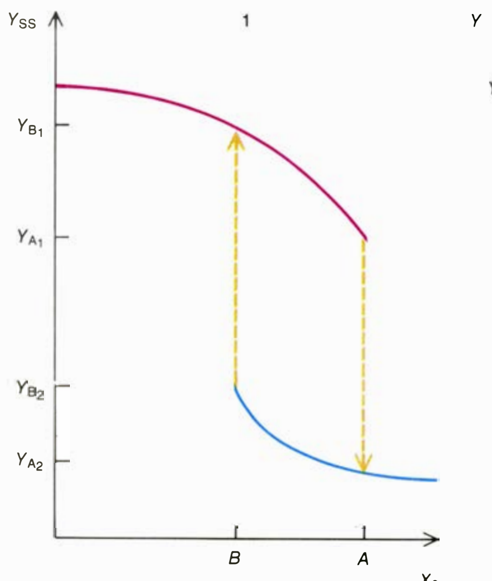
La ingente tarea de buscar el mecanismo de la reacción oscilante de BZ lo llevó a cabo, en los setenta, Richard M. Noyes, de la Universidad de Oregón, en colaboración con Richard J. Field (que se encuentre ahora en la Universidad de Montana) y Entre Körös, adscrito a la Universidad de Eötvös en Budapest. Antes de terminar 1972 habían esbozado ya un esquema que constaba de 18 pasos elementales con 20 especies químicas involucradas, capaces, tal parecía, de explicar las oscilaciones. Dos años después, una simulación detallada en el computador realizada por Noyes, Field y David Edelson, de los Laboratorios Bell, confirmaba que el mecanis-

1	$2H^+ + Br^- + BrO_3^- \rightleftharpoons HOBr + HBrO_2$
2	$H^+ + HBrO_2 + Br^- \rightleftharpoons 2HOBr$
3	$HOBr + Br^- + H^+ \rightleftharpoons Br_2 + H_2O$
4	$CH_2(COOH)_2 \rightleftharpoons (OH)_2C{=}CHCOOH$
5	$Br_2 + (OH)_2C{=}CHCOOH \rightleftharpoons H^+ + Br^- + BrCH(COOH)_2$
6	$HBrO_2 + BrO_3^- + H^+ \rightleftharpoons 2BrO_2 + H_2O$
7	$BrO_2 + Ce^{3+} + H^+ \rightleftharpoons Ce^{4+} + HBrO_2$
8	$Ce^{4+} + BrO_2 + H_2O \rightleftharpoons BrO_3^- + 2H^+ + Ce^{3+}$
9	$2HBrO_2 \rightleftharpoons HOBr + BrO_3^- + H^+$
10	$Ce^{4+} + CH_2(COOH)_2 \rightleftharpoons CH(COOH)_2 + Ce^{3+} + H^+$
11	$CH(COOH)_2 + BrCH(COOH)_2 + H_2O \rightleftharpoons Br^- + CH_2(COOH)_2 + HOC(COOH)_2 + H^+$
12	$Ce^{4+} + BrCH(COOH)_2 + H_2O \rightleftharpoons Br^- + HOC(COOH)_2 + Ce^{3+} + 2H^+$
13	$2HOC(COOH)_2 \rightleftharpoons HOCH(COOH)_2 + O{=}CHCOOH + CO_2$
14	$Ce^{4+} + HOCH(COOH)_2 \rightleftharpoons HOC(COOH)_2 + Ce^{3+} + H^+$
15	$Ce^{4+} + O{=}CHCOOH \rightleftharpoons O{=}CCOOH + Ce^{3+} + H^+$
16	$2O{=}CCOOH + H_2O \rightleftharpoons O{=}CHCOOH + HCOOH + CO_2$
17	$Br_2 + HCOOH \rightarrow 2Br^- + CO_2 + 2H^+$
18	$2CH(COOH)_2 + H_2O \rightleftharpoons CH_2(COOH)_2 + HOCH(COOH)_2$

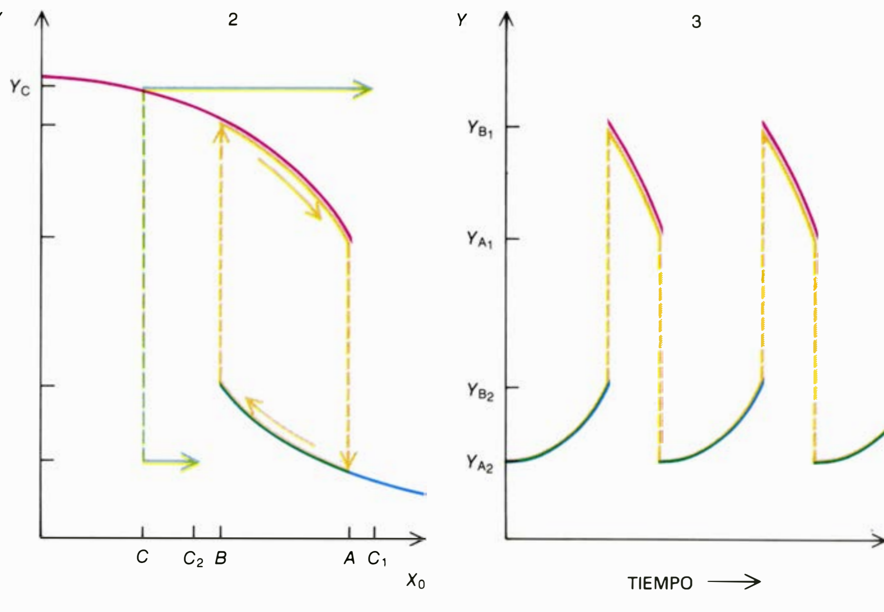
3. SIMULACION POR COMPUTADOR DE LA OSCILACION en la reacción Belousov-Zhabotinsky. Consta de 18 pasos elementales y pone en juego 21 especies químicas diferentes. La reacción lleva el nombre de dos químicos rusos: B. P. Belousov, que la descubrió, y A. M. Zhabotinsky, quien la mejoró. Se comienza con tres sustancias inorgánicas, los iones bromato (BrO_3^-), los iones bromuro (Br^-) y los iones cerosos (Ce^{3+}), y una sustancia orgánica, el ácido malónico ($CH_2(COOH)_2$). Un medio de ácido sulfúrico suministra los iones hidrógeno (H^+). Los productos de la reacción son dióxido de carbono (CO_2), ácido fórmico (HCOOH) y ácido bromomalónico ($BrCH(COOH)_2$). Como el cerio oscila entre el estado cérico (Ce^{4+}) y el ceroso (Ce^{3+}), la solución alterna entre amarilla y clara transparente. Las reacciones que se muestran arriba en color involucran solamente especies inorgánicas y se entienden mejor que las reacciones que aparecen en negro, que tratan con especies derivadas del ácido malónico. Este mecanismo lo descifraron Richard M. Noyes, de la Universidad de Oregón, Richard J. Field, su colaborador postdoctoral en el mismo centro superior, y Endre Körös, de la Universidad Eötvös.



4. REPRESENTACION DE UN SISTEMA QUIMICO BIESTABLE a través de una analogía mecánica. Una pelota que rueda en un pozo de potencial con dos mínimos *A* y *B* puede permanecer en cualquiera de ellos. *A* y *B* representan estados estables del sistema; *C*, un estado inestable. Aunque la pelota pudiera llegar a pararse en *C*, la menor perturbación la mandaría hasta *A* o hasta *B*. En este diagrama la variable es la posición de la pelota (*abscisas*); la condición externa, la forma del pozo de potencial (*ordenadas*).



5. CICLO DE HISTERESIS representado por un sistema químico biestable que opera en un reactor de flujo continuo en un tanque agitado. A medida que la concentración X_0 de la sustancia X que fluye en el reactor aumenta desde un valor bajo, el valor estable Y_{SS} de la especie Y en el reactor disminuye suavemente siguiendo la curva roja, hasta que en el punto crítico A cae de repente desde Y_{A_1} hasta Y_{A_2}. Si X_0 sigue creciendo, Y_{SS} cae de nuevo, a lo largo de la curva azul. Si se disminuye X_0, Y_{SS} torna a recorrer la curva azul y pasa el punto A hasta que alcanza un segundo punto crítico B. Aquí Y_{SS} salta desde Y_{B_2} hasta Y_{B_1}. Entre los puntos A y B se dará cualquiera de los dos estados estables.



6. SE PRODUCE UNA OSCILACION si se añade una sustancia apropiada, Z, a un sistema químico biestable que muestre histéresis. Si la sustancia Z reacciona con la sustancia Y a una velocidad baja para rendir la sustancia X, la concentración eficaz de la sustancia X_0 se modifica de forma que el sistema "ve" un valor más alto de X que el que recibe realmente del depósito que le alimenta (*izquierda*). El efecto de una cantidad dada de Z sobre X_0 (*flechas verdes*) será mayor en la rama roja del ciclo de histéresis (donde hay más Y) que en la rama azul. Supongamos que la reacción comienza con un flujo de X igual a C. Se añade ahora bastante Z, digamos la cantidad Z_0, de suerte que la X_0 eficaz sea igual a C_1 cuando el sistema esté sobre la rama roja e igual a C_2 cuando esté sobre la rama azul. En ausencia de Z, el flujo $X_0 = C$ produciría un valor estable Y_c. Con la adición de Z_0, sin embargo, el sistema sigue la curva roja a medida que el valor eficaz de X sube lentamente debido a la reacción de Y con Z. El sistema intenta alcanzar el valor de C_1, pero al llegar al valor de A cae abruptamente desde la rama roja y sufre una transición rápida a la rama azul. Allí el sistema "ve" un valor de $X_0 = C_2$, y así se desplaza hacia la izquierda siguiendo la curva azul en un intento por alcanzar el nuevo valor eficaz de X_0. Antes de conseguir este objetivo, en $X_0 = B$ salta de repente a la rama roja y se cierra el ciclo (*flechas amarillas*). X_0 tiene de nuevo el valor

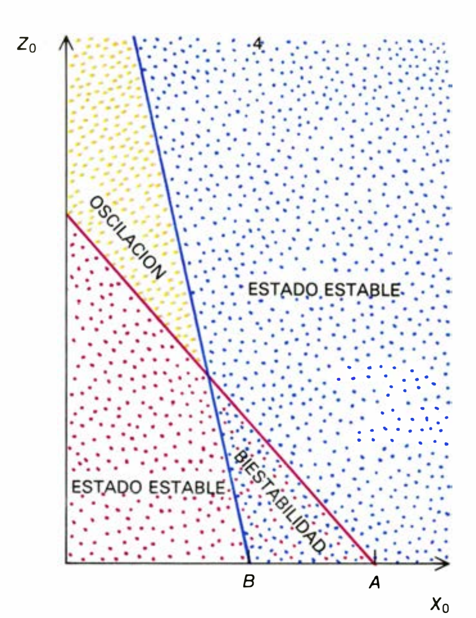
mo predecía la oscilación. Estudios posteriores han demostrado que el mecanismo justificaba también la formación de estructuras espaciales. Noyes compara su método de trabajo al detectivesco de Sherlock Holmes: "Cuando todas las demás posibilidades fallan, la que resta, *aunque improbable*, tiene que ser la cierta".

Con la solución del mecanismo de la BZ, las reacciones oscilantes alcanzaron carta de ciudadanía en química y se convirtieron en un campo apasionante para los interesados en los mecanismos y la dinámica de las reacciones. Se aventuraron varias hipótesis sobre las condiciones que se requerían para que oscilara una reacción y corrió el optimismo en algunos círculos al suponer que una mejor comprensión de las oscilaciones ayudaría a profundizar en el comportamiento periódico de los organismos vivos. Nadie ha conseguido todavía definir un conjunto de condiciones necesarias y suficientes para la oscilación química. La década de los setenta se acercaba a su fin y el número de osciladores químicos diferentes que se conocían resultaba descorazonador por lo exiguo. Todos se habían descubierto por casualidad; sólo se había descifrado uno en sus pormenores; nadie sabía qué hacer para encontrar más.

Los sistemas biológicos constituyen una buena fuente de preparaciones donde aparecen oscilaciones, pero desafían el análisis mecanístico que conduce a la generalización. Podemos preparar estos sistemas oscilantes extrayendo el contenido químico de las células de levadura y separándolo de las paredes celulares y otras estructuras extrañas. Si agregamos los nutrientes adecuados a la "sopa" resultante, la concentración de los iones hidrógeno y de la molécula nicotinamidin adenín dinucleótido, que en su forma reducida (NADH) transporta electrones en el ciclo energético normal de la célula, oscilarán. En la Johnson Research Foundation, Britton Chance y sus colaboradores han demostrado que muchos de estos osciladores biológicos recurren a la glicólisis, proceso mediante el cual las células generan energía metabolizando el azúcar en ausencia de oxígeno. Por hallarse dichas reacciones catalizadas por una serie de enzimas, presentan una complejidad inabordable.

En 1979 dos de nosotros (Epstein y Kustin) pensamos que había llegado el momento de inventar un procedimiento sistemático para diseñar un oscilador químico. ¿Por dónde comenzar? Por identificar tres condiciones que sabíamos eran necesarias para la oscilación química o conducían a ella. Primera condición: los sistemas químicos oscilan sólo si están lejos del equilibrio. La segunda es la de retroalimentación o feedback: un producto de un paso de la secuencia de reacciones ha de influir en su propia velocidad de formación. La tercera condición decía que el sistema químico debía presentar biestabilidad. Bajo el mismo conjunto de condiciones externas, a las que nos referimos como factores limitantes, el sistema ha de poder existir en dos estados estables diferentes.

La exigencia de que el sistema se mantenga lejos del equilibrio se lograba desarrollando la reacción en un aparato que conocen bien los ingenieros químicos: el reactor de flujo continuo en un tanque agitado (CSTR, del inglés "continuous-flow stirred-tank reactor"). Mediada la década de los setenta, Adolphe Pacault y sus colaboradores, en el Centro de Investigación Paul Pascal de Burdeos, habían adaptado el CSTR para el estudio de reacciones químicas oscilantes. Cuando uno de los firmantes (De Kepper) abandonó Burdeos, en 1980, para unirse a nuestro grupo de la Universidad de Brandeis, se vio lógico construir un reactor similar. La llegada de Orbán desde Hungría, experto en osciladores bromato, completó el grupo.



eficaz de C_1, de forma que el sistema se mueve otra vez hacia la derecha. La repetición de esta secuencia resulta en una oscilación periódica en el valor de Y (*en el medio*). El diagrama de fases en forma de cruz (*ilustración de esta misma página*) muestra como aumentando el valor de Z_0 se puede desplazar el sistema de la biestabilidad a la oscilación. Si Z se hace 0, el sistema exhibe entonces dos condiciones estables; a saber: el estado rojo para un X_0 pequeño y el estado azul para un X_0 grande. Cuando X_0 se encuentra entre los puntos A y B, el sistema es biestable. A medida que Z_0 aumenta, la región de biestabilidad se estrecha hasta que, a un determinado valor crítico, comienza la oscilación.

Una clase común de retroalimentación ("feedback") que se encuentra en los organismos es la autocatálisis; en ella, la tasa de producción de una sustancia aumenta con su concentración. La idea se generaliza a poblaciones de organismos: el número de organismos en que crece una población en un tiempo dado es casi siempre proporcional al número de individuos ya existentes en la población. (Esta observación reitera la predicción de Thomas Malthus en 1978, según la cual las poblaciones tienden a crecer más deprisa que los alimentos.) En química, los sistemas autocatalíticos son raros, pero no extraordinarios. En 1920 Alfred J. Lotka, de la Universidad de Johns Hopkins, demostraba que un esquema sencillo de dos reacciones autocatalíticas acopladas provocaría la oscilación. Aunque no se ha conseguido aplicar a ninguna reacción química, el mecanismo de Lotka ha guiado el pensamiento de muchos químicos y ha servido de mucho a la hora de explicar las oscilaciones de poblaciones depredadoras y sus presas en los ecosistemas.

Vayamos con la tercera condición: la biestabilidad, o existencia de dos estados estables diferentes, en un sistema químico oscilante. Por estado estable se entiende el sistema en que todas sus variables, así la concentración de cada especie química, han alcanzado valores constantes. De un estado decimos que es estable si es capaz de acomodarse a un pequeño cambio en una variable, la adición de una gota de ácido por ejemplo, y se recupera sin transformarse en otro nuevo estado. Si el menor cambio causa una transición a un estado diferente del sistema, se dice que es inestable. Por ejemplo, una canica en el fondo de un cuenco se halla en un estado estable; ella misma balanceándose en el borde, se encontrará en un estado inestable. En un reactor de flujo continuo y agitado, los factores limitantes o parámetros externos que controlan los valores estacionarios del sistema reaccionante suelen ser: la temperatura del baño termostático, la concentración de las sustancias químicas en el depósito que alimenta el reactor y sus velocidades de flujo.

Si alteramos los parámetros externos de un sistema químico biestable, puede presentarse un fenómeno peculiar llamado histéresis. Se trata de un fenómeno corriente en magnetismo, donde adquiere la forma de "ciclo de histéresis". Cuando sometemos un trozo de hierro a un campo magnético cada vez más intenso, el hierro termina por alcanzar su completa saturación magnética. Si se desconecta el campo, tomando de nuevo el valor cero, el hierro retiene parte de su magnetización. Para reducir la magnetización del hierro de nuevo hasta cero, es necesario someter la muestra a un campo de sentido opuesto. El ciclo entero de magnetización y desmagnetización toma la forma de un bucle en forma de S, ancho en el centro y puntiagudo en los extremos.

Un fenómeno análogo puede ocurrir en un sistema químico biestable. En 1979, Jacques Boissonade y uno de los autores (De Kepper) realizaron, en el Centro de Investigación Paul Pascal, un estudio del que se desprendía que las oscilaciones podían inducirse al añadir otra sustancia química a un sistema capaz de histéresis. El comportamiento de este sistema se representaba en un diagrama de fases en forma de cruz, donde se evidenciaba que si la sustancia añadida ejercía efectos significativamente diferentes sobre las dos ramas de los estados estables del sistema, era probable que surgieran oscilaciones.

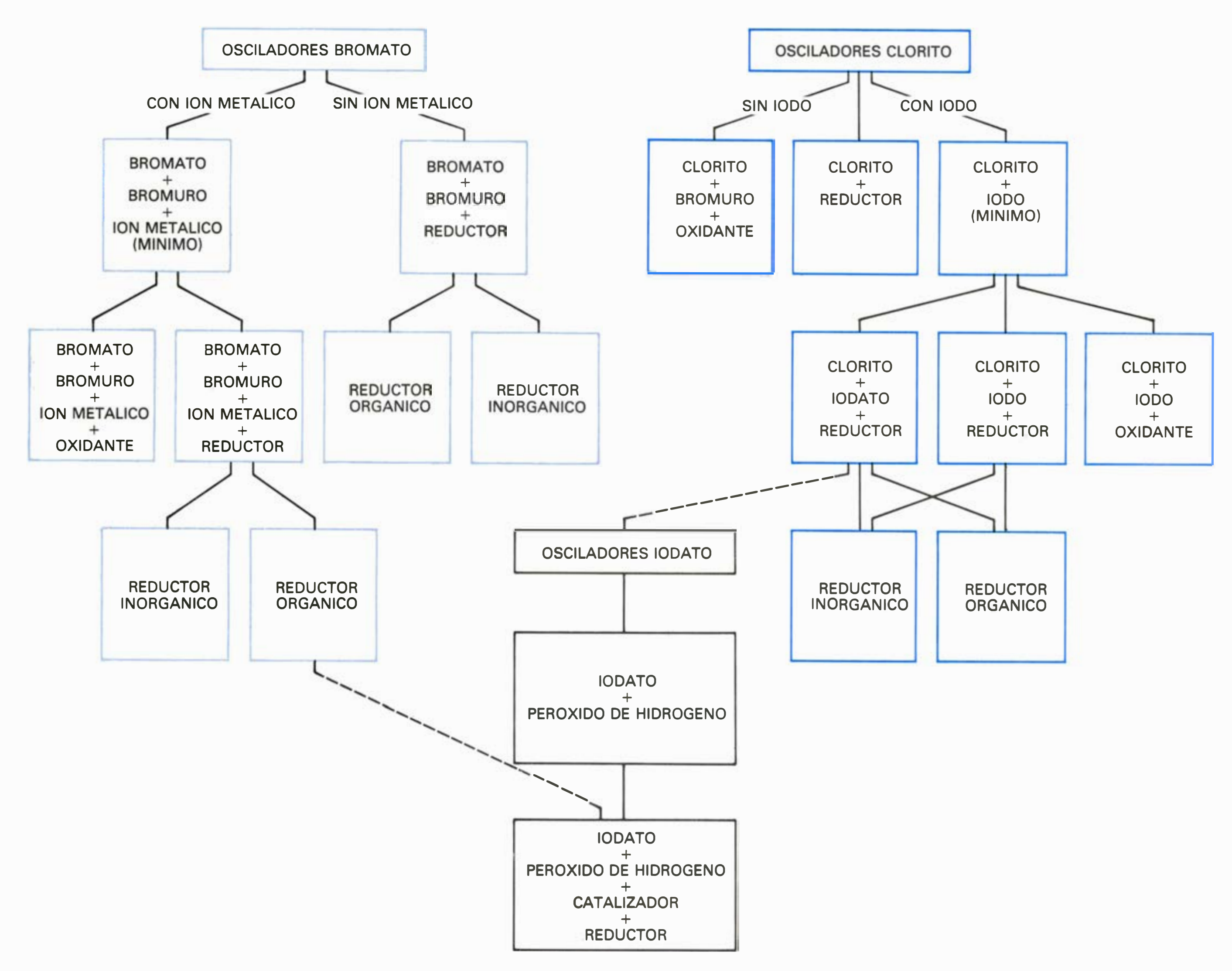
Nuestro plan de trabajo para diseñar un oscilador químico comprende los cuatro pasos siguientes: encontrar un sistema autocatalítico, llevar a cabo la reacción en un reactor de flujo continuo, variar las condiciones hasta que se encuentre una región de biestabilidad y entonces introducir otra sustancia que

sea capaz de afectar las dos ramas de biestabilidad de forma diferente e inducir así oscilaciones. Buscamos bibliografía sobre reacciones autocatalíticas; dimos con algunas que prometían. Había dos en particular muy sugestivas, ya que compartían el iodo como especie intermedia común; una involucraba iones iodato y arsenito y, la segunda, iones ioduro y clorito. ¿Podía tener una reacción el efecto perturbador deseado sobre la otra? Pronto se comprobó que la reacción iodato-arsenito era biestable. Introdujimos clorito y el sistema compuesto comenzó a oscilar, casi inmediatamente. El primer empeño sistemático aplicado al diseño de un nuevo oscilador químico resultó un éxito.

Variando los componentes que entran en el reactor de flujo continuo, hemos ampliado desde entonces el oscilador clorito-iodato-arsenito del comienzo hasta reunir una familia numerosa de dos docenas de sistemas relacionados que tienen el ion clorito en común. Podemos ya clasificar una taxonomía de los osciladores clorito. Habrá que relacionarla con las familias de los osciladores de iodatos y bromatos; y en eso estamos.

Aunque los osciladores bromato se cuentan entre los primeros en haberse descubierto, y se comprende mejor su mecanismo que el de los osciladores clorito, la gama de sistemas clorito que se han encontrado a través de un programa de trabajo es mucho más amplia. Dos factores, uno químico y otro histórico, justifican esta paradoja. Químicamente, cada familia de osciladores posee un miembro unidad, o más sencillo, a partir del cual se pueden deducir otros añadiendo nuevas sustancias. El oscilador clorito unidad, formado por clorito más ioduro, oscila en un intervalo de condiciones amplias. Por contra, la unidad del oscilador bromato, bromato más bromuro más un ion metálico, opera sólo bajo condiciones muy precisas. La verdad es que ni siquiera después de que Kedma Bar-Eli, de la Universidad de Tel-Aviv, predijera su existencia la búsqueda del oscilador bromato unidad dio resultados hasta que dos de los autores (Orbán y Epstein) lograron hallarlo tras buscarlo sistemáticamente a través del diagrama de fases de forma cruzada.

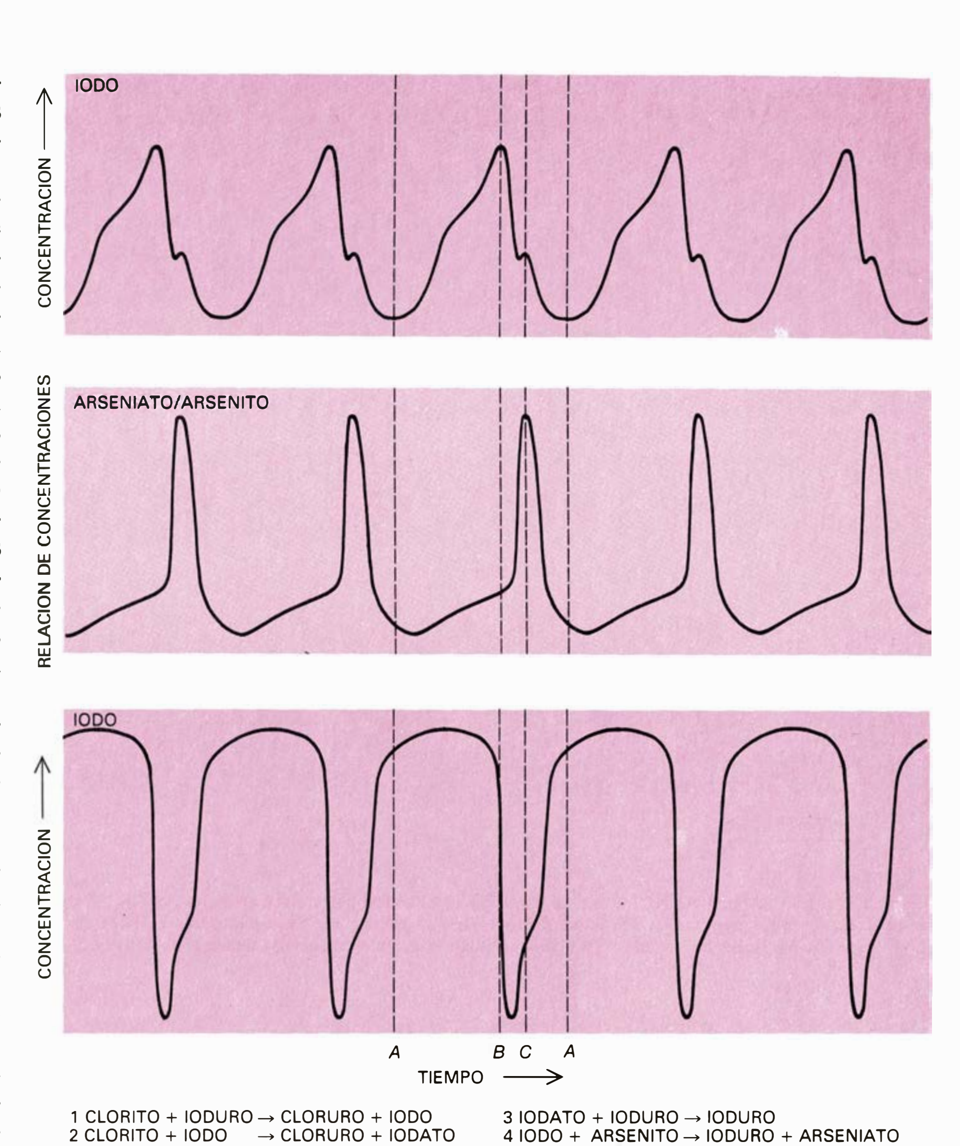
Desde el punto de vista histórico, los primeros osciladores bromato se descubrieron antes de que se utilizaran los reactores de flujo continuo. Los sistemas bromato oscilaban fácilmente en condiciones de flujo nulo. Los osciladores clorito, como la gran mayoría de los



7. PRINCIPALES CLASES DE OSCILADORES QUIMICOS, agrupados en la ilustración, según una taxonomía provisional. Los hemos reunido en tres familias principales. El oscilador Belousov-Zhabotinsky y sus muchas variantes se encuentra en la rama de la familia del bromo que tiene iones metálicos; el oscilador bromato en la rama que está libre de iones metálicos constituye el trabajo de Körös y uno de los autores (Orbán) en la Universidad de Eötvös. Todos los osciladores clorito han sido descubiertos por los autores en Brandeis. Los sistemas bromato que implican reductores inorgánicos fueron desarrollados en Brandeis en condiciones de flujo continuo. La familia iodato consta de menos miembros. El primer sistema que se menciona en este grupo es la reacción oscilante que estudió en 1921 William C. Bray, de la Universidad de California en Berkeley, y que permaneció ignorada durante 50 años. El segundo sistema de la familia del iodato es un híbrido de los sistemas Belousov-Zhabotinsky y de William C. Bray. Las líneas a trazos conectan sistemas que tienen ciertos reactivos en común. Una sustancia reductora es aquella que dona electrones a otra sustancia; la oxidante capta electrones.

nuevos sistemas, muestran un comportamiento periódico sólo en condiciones de flujo continuo, porque el flujo es necesario para mantenerlos lejos del equilibrio. Un sistema que oscile en condiciones normales, sin flujo, casi siempre lo hará en condiciones de flujo continuo; lo contrario ocurre muy pocas veces. La mayoría de los osciladores bromato se buscaron originalmente en condiciones normales, en tanto que los osciladores clorito se buscaron sólo en condiciones de flujo. De ahí que los múltiples sistemas clorito que se descubrieron pronto sobrepasaran el número de sistemas bromato. (Los autores hallarían más tarde que varios osciladores clorito operaban también en condiciones normales.) Podría decirse que la facilidad con que oscilaron los primeros sistemas bromato bloqueó la búsqueda de otros nuevos. Sólo recientemente se han utilizado los reactores de flujo continuo para identificar nuevos osciladores bromato y estudiar la tercera familia de osciladores: los sistemas iodato; hasta hoy, la familia más exigua.

Al diseñar cada nuevo oscilador químico, nos proponíamos obtener un sistema cuyo mecanismo pudiera descifrarse y compararse con los de las reacciones oscilantes ya examinadas antes. La disección de la familia de los sistemas oscilantes clorito lleva buen camino. Se han estudiado las velocidades de varios procesos que la componen y empieza a haber varias piezas del rompecabezas; la descripción completa del esquema oscilante parece ser cuestión de ir acomodando los grupos hasta que casen todos.

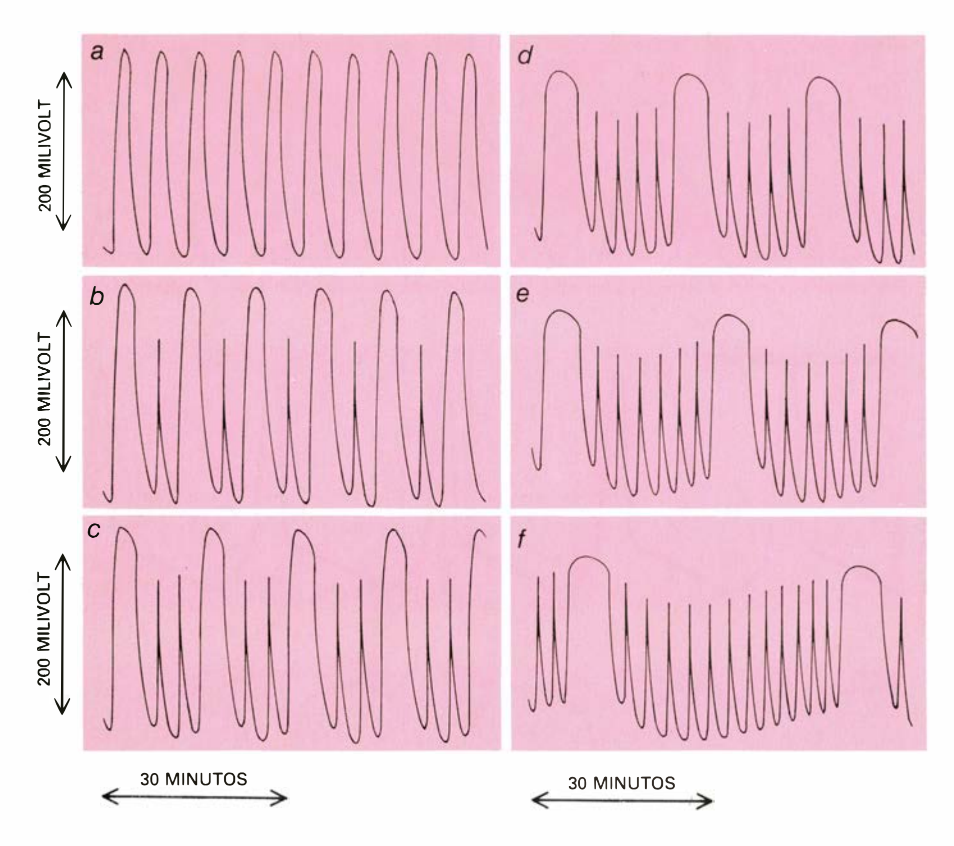


8. PAUTA DE OSCILACION observada en el primer sistema oscilante ideado por los autores. Puede interpretarse cualitativamente a partir de la interacción entre cuatro reacciones principales. El sistema oscilante se mantiene a través de un flujo, uniforme y equilibrado, de tres soluciones salinas: un clorito, un arsenito y un iodato. En el seno del reactor de flujo continuo, las concentraciones de iodo y iodato, así como la razón de arsenato a arsenito, oscilan, según nos manifiestan las tres curvas. En el punto *A* están subidos el ioduro y el arsenito, en tanto que iodo y arseniato se hallan bajos; por tanto, predominan las reacciones *1* y *3*. Estas reacciones aumentan drásticamente la concentración de iodo e inician el descenso de la concentración de ioduro, llevando el sistema hasta el punto *B* y suspendiendo las reacciones *1* y *3*. Intervienen entonces las reacciones *2* y *4*, rebajando el iodo y convirtiendo el arsenito en arsenato. El sistema alcanza el punto *C*, donde el arsenito y el iodo se encuentran en niveles mínimos, o en su vecindad. La alimentación del reactor hace que se recupere el arsenito mientras baja el arsenato por salida del mismo. La reacción *4* regenera el ioduro al tiempo que consume iodo, devolviendo el sistema al punto *A*.

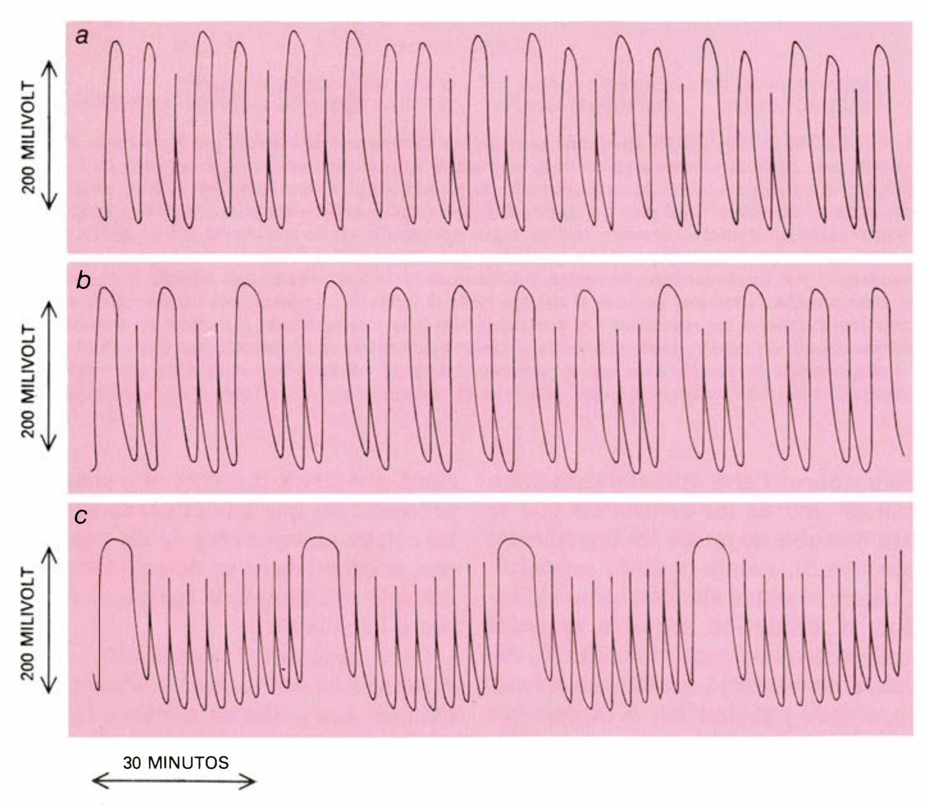
Se puede dar ya una visión esquemática de cómo funciona nuestro primer sistema oscilante, en donde entran soluciones de sodio o potasio en forma de clorito, iodato y arsenito. Consta de sólo cuatro reacciones fundamentales; cada una se compone de varios pasos elementales. Dos de ellos implican la reacción del clorito con el ioduro o con el iodo, rindiendo cloruro e iodo en el primer caso y cloruro e iodato en el segundo. En la tercera reacción, el iodato reacciona con el ioduro para dar iodo. En la cuarta, iodo y arsenito forman ioduro y arseniato. Las variaciones en las velocidades de estas cuatro reacciones originan oscilaciones periódicas en las concentraciones de iodo e iodato y en la razón de arseniato a arsenito [*véase la figura 8*].

La oscilación química, fascinante en sí misma, está relacionada con otros fenómenos de no menor interés; por ejemplo, con la formación de estructuras espaciales en un medio inicialmente homogéneo. Estas estructuras se originan en uno de los osciladores que se han descubierto donde los ingredientes son clorito, ioduro y ácido malónico. Cuando se añade almidón como indicador, la disolución adquiere un color azul oscuro uniforme al comienzo, debido a un complejo que forman el ioduro, el iodo y el almidón. A medida que avanza la reacción, aparecen puntos blancos que se desarrollan en anillos y conjuntos de anillos concéntricos, que se aniquilan entre sí al chocar. Alguien comparó la aparición de los puntos blancos contra el fondo azul oscuro con la aparición de las estrellas al atardecer. La presencia de estas formas geométricas en una masa antes indiferenciada nos trae a la mente otra clase de procesos: los que tal vez permiten que las células embrionarias de los organismos se especialicen en determinados tipos que constituirán la sangre, el cerebro o los huesos.

Otro fenómeno relacionado, cuya existencia ha desatado en la comunidad química una polémica similar a la que suscitó hace un cuarto de siglo la reacción oscilante, es el del "caos químico". La expresión alude a una reacción en que las concentraciones ni alcanzan valores constantes ni oscilan periódicamente, sino que aumentan y disminuyen al azar y de forma imprevisible. Hemos observado estas fluctuaciones en el sistema BZ y en uno de nuestros



9. OSCILACIONES PERIODICAS en la reacción clorito-tiosulfato. Muestran un amplio abanico de comportamientos cuando la velocidad de inyección al reactor de flujo continuo se altera de un modo sistemático. Cada ciclo tiene una gran oscilación y otras *n* pequeñas, donde *n* varía de 0 a 16.



10. OSCILACION APERIODICA, o caótica. Se observa a veces en un intervalo estrecho de velocidades de flujo entre velocidades que dan origen a una conducta periódica compleja. Tenemos un tipo de caos en una mezcla aparentemente al azar de dos clases de oscilación periódica, una con *n* y otra con *n* + 1 oscilaciones pequeñas por cada oscilación grande. Podría explicarse diciendo que variaciones no controladas en la velocidad de flujo modifican el sistema al azar adelante y atrás, desde una velocidad que da una oscilación de la clase *n* hasta otra que da la clase (*n* + 1). La alternativa, que defienden los autores, afirma que la oscilación caótica es intrínseca de la dinámica de la reacción química a ciertas velocidades.

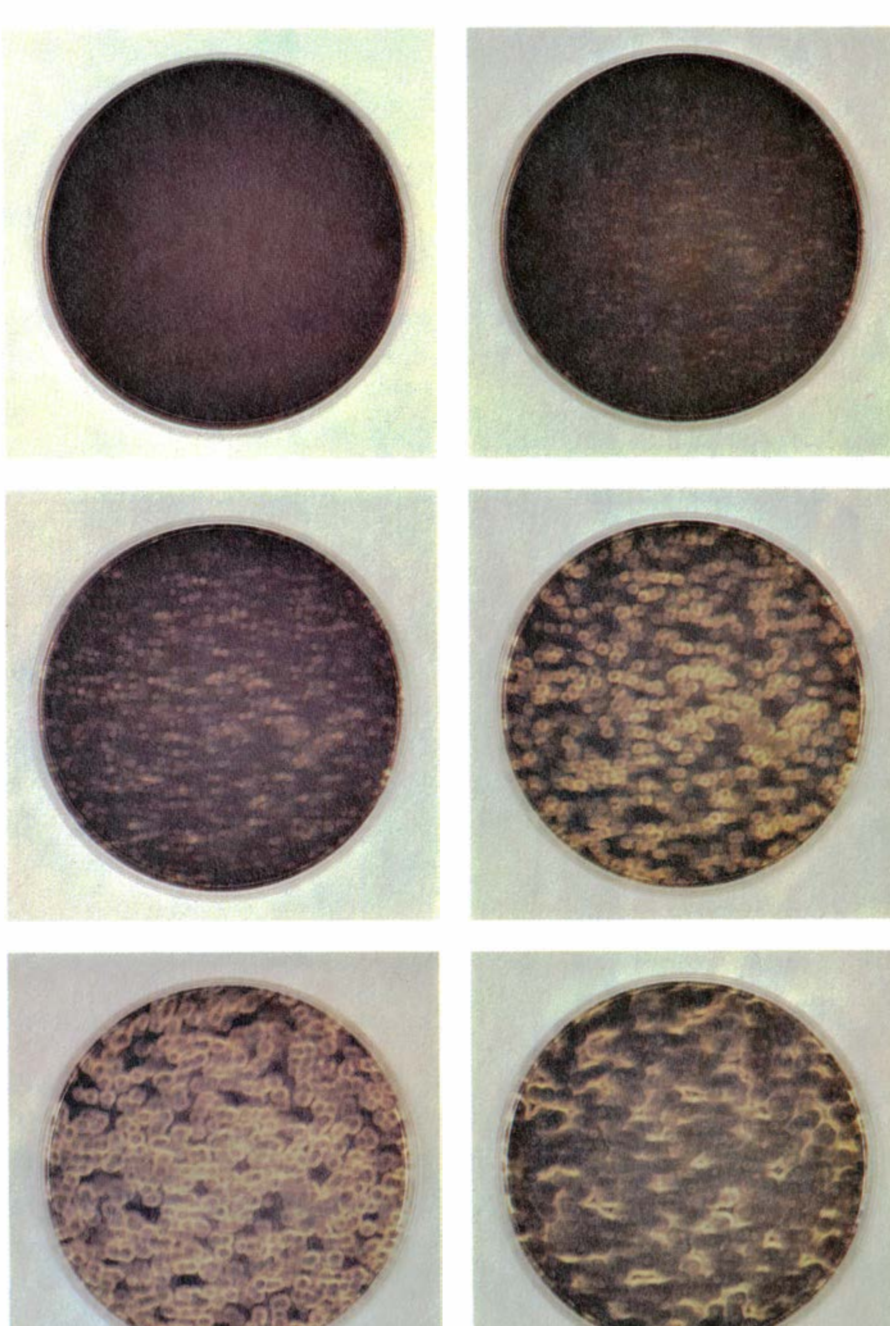
nuevos osciladores clorito, cuando se les hacía reaccionar en un reactor de flujo continuo sometido a riguroso control. Presentan un interés especial para los matemáticos: si fueran parte genuina de la dinámica de la reacción, constituirían ejemplos físicos de ciertos conceptos matemáticos, los llamados atractores extraños. Dichos atractores extraños surgen cuando se iteran determinadas ecuaciones, o se resuelven repetidamente, con el resultado de que una iteración sirve de información ("input") para la siguiente. En un sistema de ecuaciones con un atractor extraño el camino que describen las sucesivas soluciones varía de un ciclo al siguiente de forma impredecible. (Douglas R. Hofstadter lo exploró en la sección de "Temas metamágicos" de enero de 1982.)

Los químicos escépticos atribuyen el caos químico a posibles fluctuaciones experimentales incontroladas, en la temperatura o en la velocidad de flujo por ejemplo, lo que desplazaría al azar el sistema de un modo de oscilación periódica y compleja a otro. Aunque la pelota está aún en el tejado, trabajos experimentales y teóricos realizados con extrema finura por Harry L. Swinney y Jack Turner, de la Universidad de Texas en Austin, Jean Claude Roux, Christian Vidal y su grupo, del Centro de Investigación Paul Pascal, y por J. L. Hudson, de la Universidad de Virginia, sugieren firmemente que el caos químico es un fenómeno genuino de la dinámica de al menos algunos osciladores químicos. No sabemos todavía dónde nos llevarán las implicaciones de un caos intrínseco. ¿Por qué no pensar, como afirmara en cierta ocasión Henry Adams, que "el caos engendra vida, cuando el orden crea hábito"?

Demostrada la coherencia de los osciladores químicos con las leyes de la naturaleza y nuestra posibilidad de diseñarlos y describirlos sistemáticamente, comienzan a hacer acto de presencia, con un papel destacado, en varias áreas de la ciencia. El cambio de un estado a otro, distintivo de la reacción oscilante, quizá guarde la clave que nos descifre algunos procesos reguladores de la célula: los mecanismos que inician o bloquean la transcripción de una cadena de ADN o la contracción de un músculo. Las mismas fuerzas que crean anillos coloreados y capas en osciladores químicos sin agitar podrían tener que ver con las responsbles de las distancias que median entre los anillos de Saturno y las estrías periódicas que se encuentran en ciertas forma-

ciones rocosas, que no encuentran una explicación cómoda a través de los procesos geológicos convencionales.

También parece plausible que muchas, si no todas, las reacciones catalíticas que son fundamentales en la industria química tengan lugar de forma oscilatoria. El catalizador alternaría, en este supuesto, entre dos formas a lo largo de un intervalo temporal demasiado corto para detectarlas por los métodos de análisis clásicos. El estudio de las reacciones oscilantes se convertirá en banco donde probar las teorías de catálisis, y de modo más general, las teorías que describen la dinámica de las reacciones químicas. Cualquier modelo que prediga con éxito la compleja conducta de las reacciones oscilantes debe interpretar fácilmente los fenómenos más sencillos que se observan en reacciones menos exóticas. A medida que se diseñen o se encuentren nuevos osciladores químicos, se irá recogiendo información para abordar, desde una nueva perspectiva, múltiples fenómenos, incluidos los relojes y toboganes que operan dentro de nosotros.



11. SE DESARROLLA UNA ESTRUCTURA EN TRES DIMENSIONES en un oscilador basado en el clorito, ioduro y ácido malónico. La solución, colocada en un plato llano, arranca desde un morado uniforme (*arriba, a la izquierda*) de un complejo de iodo, ioduro y un indicador de almidón. A medida que la reacción avanza aparecen puntos blancos; éstos crecen formando anillos y conjuntos de anillos concéntricos, que se aniquilan entre sí al chocar. La última foto se tomó a los 90 segundos de empezar la reacción.

Halobacterias

Crecen a concentraciones saturantes de sales, transportan iones utilizando la energía lumínica y modifican continuamente su material genético

Francisco Rodríguez Valera y Francisco Ruiz Berraquero

A principios de siglo, científicos de diversas partes del mundo se interesaron por un fenómeno que producía preocupación y miedo a los almacenistas y distribuidores de salazones de pescado: el "mal rojo del bacalao". Cuando una pieza se veía afectada, aparecía una capa de color rojo intenso que cubría en todo o en parte su superficie confiriéndole un aspecto muy poco apetitoso, unido a un ligero olor a descomposición que la desechaban para el consumo. Pronto se comprobó que todo el cuadro se debía a la presencia de bacterias con un pigmento rojo muy llamativo que proliferaban a concentraciones saturantes de sal, siempre que la humedad y la temperatura fuesen elevadas. Hoy sabemos que los microorganismos que originan el "mal rojo del bacalao" pertenecen a la familia de las Halobacteriáceas, que incluye dos géneros: uno formado por bacilos (género *Halobacterium*) y otro por cocos (género *Halococcus*). No se trata de bacterias típicas, sino de archibacterias: microorganismos que parecen estar relacionados tanto con las bacterias clásicas como con los organismos eucariotas [véase "Archibacterias", por Carl R. Woese; INVESTIGACIÓN Y CIENCIA, agosto de 1981].

El estudio de la fisiología de las halobacterias nos ha revelado unos mecanismos fascinantes que ensanchan de forma sorprendente la plasticidad de la vida. De hecho, y considerando las grandes lagunas que existen en el conocimiento de estos organismos, parece muy probable que todavía nos deparen nuevas sorpresas. Todas las archibacterias viven actualmente en medios que pueden considerarse extremos, al menos desde el punto de vista de las condiciones que han imperado en la biosfera durante los últimos dos mil millones de años. Es posible que antes, durante las primeras etapas de la evolución de la vida, las archibacterias constituyesen un grupo ampliamente distribuido, que quedó arrinconado ante la expansión de grupos más modernos y mejores competidores en hábitats poco deseables: los metanógenos en ambientes completamente exentos de oxígeno, los termoacidófilos en manantiales termales y las halobacterias en medios hipersalinos. Decimos que un medio es hipersalino cuando su concentración de sal es superior a la del agua de mar. ¿Escasean los medios hipersalinos? Tenemos bastantes pruebas de que los medios de este tipo han sido muy abundantes y han aparecido con frecuencia en el tiempo y en el espacio. El origen primario de su existencia es el mar. El mar contiene cantidades considerables de sales, siendo el sodio (Na^+), magnesio (Mg^{2+}), calcio (Ca^{2+}) y potasio (K^+) sus principales cationes y cloruros (Cl^-), sulfatos ($SO_4^=$), carbonatos (HCO_3^-) y bromuros (Br^-) los aniones más importantes. La evaporación del agua de mar permite la concentración de estas sales que, además, mantienen sus proporciones relativas iguales a las que presentan en el mar, puesto que, a excepción del Ca^{2+} que precipita como sulfatos y carbonatos en los primeros estadios de concentración, las sales formadas por los restantes son todas muy solubles.

Este proceso de evaporación es el que se produce en las salinas solares construidas por el hombre para obtener sal común. Las salinas solares mimetizan un tipo de estructura natural que podríamos llamar salinas naturales. En las salinas naturales, el agua salada percola hasta zonas deprimidas junto al mar, formando lagunas más o menos extensas donde se concentran las sales. Estos medios abundan en las costas bajas de lugares del mundo muy distantes, como son las costas mediterráneas y las de la Baja California; originan un tipo característico de roca sedimentaria denominado evaporitas, rocas formadas por depósitos de sales marinas que existieron en un pasado remoto. La sal de roca tiene también un origen semejante. Podemos decir que las salinas naturales son tan antiguas como el propio mar. Probablemente siempre hubo más que hoy, cuando el hombre las deseca para cultivar el terreno o construir complejos residenciales junto al mar.

Como hábitat, las salinas, amén de la elevada concentración de sal, presentan otras características. Al ser aguas someras situadas en zonas cálidas y secas, están sometidas a una fuerte irradiación solar. Las temperaturas son elevadas y oscilan fuertemente del día a la noche. Puesto que el agua con altas concentraciones de sales disueltas tiene un calor específico bajo, se calentará muy deprisa por las mañanas alcanzando fácilmente 45 grados o 50 grados C en un día caluroso de verano, descendiendo durante la noche hasta casi igualar la temperatura del aire. No es pues un lugar confortable para organismos que buscan la regularidad de las condiciones de su ambiente. Otra característica de los medios hipersalinos es su bajo contenido en oxígeno, cuya solubilidad en esos medios tan concentrados es prácticamente nula. Sin embargo, no todo son condiciones negativas. La evaporación del agua de mar, además de concentrar sales, concentra también los nutrientes orgánicos e inorgánicos, proporcionando a los organismos halófilos un medio bastante rico.

Algunos hábitats hipersalinos no se originan en el mar. Son los llamados lagos salados, cuya formación se debe a la disolución de depósitos minerales de sales en proporciones que a veces se alejan de las que encontramos en un medio marino típico; tal es el caso del mar Muerto, cuyo contenido en magnesio es muy elevado, superando incluso al de sodio. Desde el punto de vista teórico, estos medios son menos interesantes y no se les puede considerar el hábitat primario de los organismos halófilos, por darse de una forma esporá-

dica, sin la frecuencia de las salinas. El mar Muerto tiene escasos miles de años de vida y es muy probable que nunca haya existido otro hábitat similar. Hace siete años, los autores, trabajando en el departamento de biología de la Universidad de Alicante, descubrieron unas salinas situadas junto a la población de Santa Pola (Alicante), que en muchos aspectos constituían un magnífico modelo de salina natural o, más exactamente, de numerosas salinas naturales. En estas salinas, como en muchas de las existentes en la costa mediterránea, la sal se obtiene mediante un proceso continuo: el agua de mar circula a través de un sistema de estanques donde se va evaporando y en consecuencia se origina un gradiente de concentraciones. En cada estanque, la concentración de sales se mantiene aproximadamente constante en el tiempo, lo que supone disponer de decenas de salinas naturales, cada una de ellas con una concentración de sales diferente: desde ligeramente superior al agua de mar hasta un 50 por ciento en sales. Si un investigador que quisiera estudiar los organismos halófilos y su ecología pudiera diseñar un sistema artificial donde poder estudiar la gama de hábitats hipersalinos que se originan en la evaporación del agua de mar, probablemente diseñaría un sistema de este tipo.

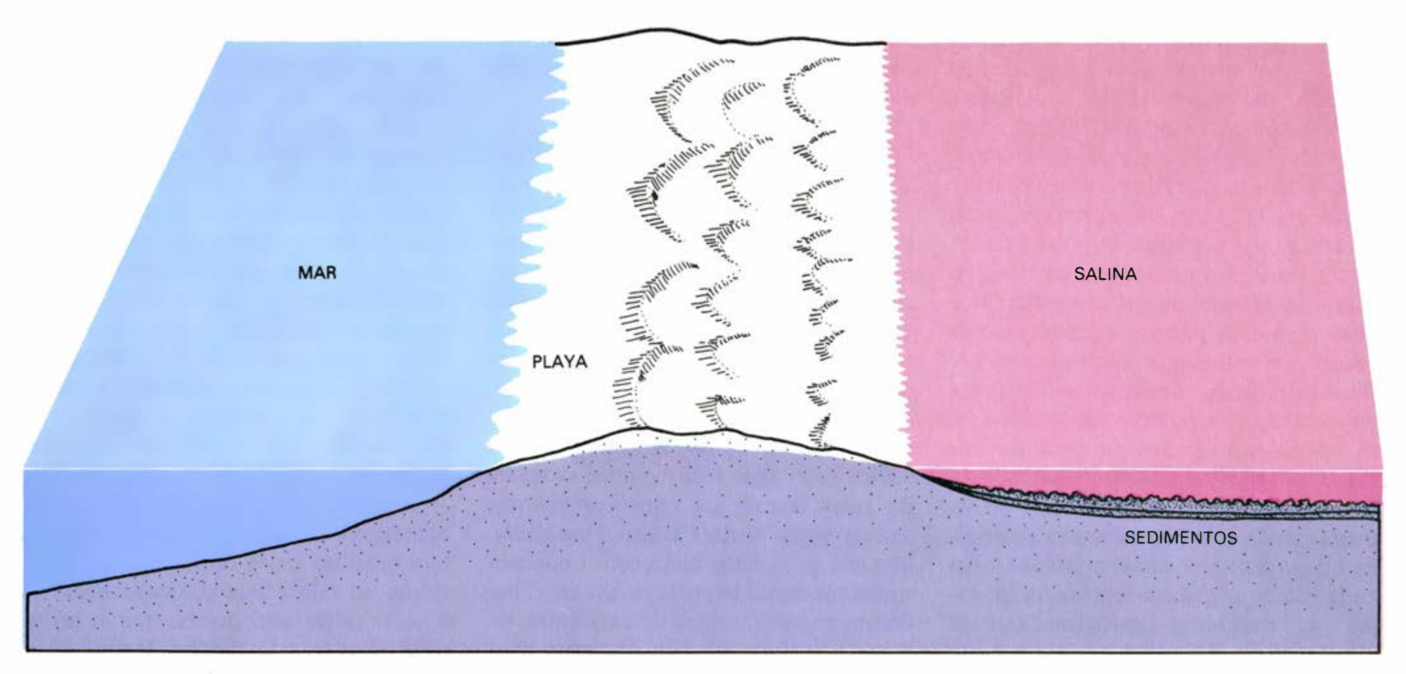
Aunque las halobacterias son indudablemente los habitantes más conocidos de los medios hipersalinos, no son los únicos. A pesar de la aparente hostilidad del ambiente, es notable la variedad de los organismos que viven allí: organismos pluricelulares como el crustáceo *Artemia salina* y las larvas de la mosca *Ephydra*; plantas superiores como algunos *Arthrocnemum* que se desarrollan con sus raíces sumergidas en suelos cuya agua está siempre saturada de sales; entre los microorganismos, vemos algas verdes como *Dunaliella*, algunos protozoos aparecen a concentraciones de sal bajas y, finalmente, hallamos representantes de casi todos los grupos bacterianos que se encuentran en el mar, incluyendo cianobacterias, bacterias fotosintéticas y quimiorganotrofos aerobios y anaerobios. (Los quimiorganotrofos designan a los que utilizan compuestos orgánicos del medio como fuente de energía.) La diversidad disminuye con la concentración de sales y llega un momento en que sólo medran las halobacterias.

¿Son halófilos todos los organismos que viven en medios hipersalinos? Un organismo se considerará halófilo cuando, para su desarrollo, requiere concentraciones de sal más altas que la del agua de mar. Los organismos que pueden desarrollarse a concentraciones altas de cloruro sódico, pero que crecen mejor sin sales, no son halófilos sino halotolerantes, ya que no están específicamente adaptados a vivir en medios hipersalinos; solamente los toleran. En contraste, el auténtico halófilo ha experimentado un proceso evolutivo de adaptación y ha modificado irreversiblemente su fisiología en estos medios. Aunque la halotolerancia de algunas bacterias es sorprendentemente alta (se han descrito casos de hasta un 20 por ciento en sales) hemos comprobado en nuestra salina que, por encima de un 10 por ciento en sales totales, la población bacteriana es masivamente halofílica con muy pocos representantes halotolerantes. Y tal puede ser el probable cómputo de los demás grupos.

¿Qué cambios sufrieron en su fisiología los organismos halófilos para convertirse en tales? Aquí nos vamos a centrar en los microorganismos, puesto que los pluricelulares, con sus diferen-



1. VARIABILIDAD ESPONTANEA DE *Halobacterium halobium*, un halófilo extremo que forma vacuolas de gas (órganos de flotación), que dan a las colonias un aspecto opaco. En las poblaciones de esta bacteria hay individuos que no sintetizan vacuolas y forman colonias transparentes. Esta variabilidad fenotípica se debe a la inserción de cortos segmentos de ADN, transposones, en un plásmido. Se trata de un fenómeno frecuente en el genoma de las halobacterias, hasta el punto de habérsele bautizado con el nombre de genoma elástico. Refleja una organización peculiar, genuina de las archibacterias.



2. FORMACION DE UNA SALINA NATURAL. El agua de mar percola en zonas deprimidas junto a la costa, donde la evaporación concentra las sales hasta alcanzar niveles de saturación. Las lluvias pueden rebajar la concentración de sales; se establece, por tanto, una cierta periodicidad: concentraciones muy altas en épocas muy cálidas y secas que bajan en épocas lluviosas y frías. Estas salinas naturales se originan, con mucha frecuencia, en costas bajas.

tes tejidos, pueden aislarse del medio externo y mantener un medio interno que nada tiene de hipersalino. Para ellos, el problema se limita a bombear suficiente agua del exterior; ésta es la razón de que las plantas halófilas sean muy parecidas a simples xerófitos.

Los microorganismos no pueden aislarse. El primer problema que se les plantea será el osmótico. Todas las membranas plasmáticas son permeables al agua; por ello, cuando la concentración de solutos dentro de la célula es mayor que en el exterior (como ocurre casi siempre), el agua del exterior tiende a entrar en la célula creando una presión positiva que haría estallar la célula, si no fuera por su pared rígida que contrarresta dicha presión. Si bien este problema es de fácil solución, no lo es el opuesto: ante una concentración



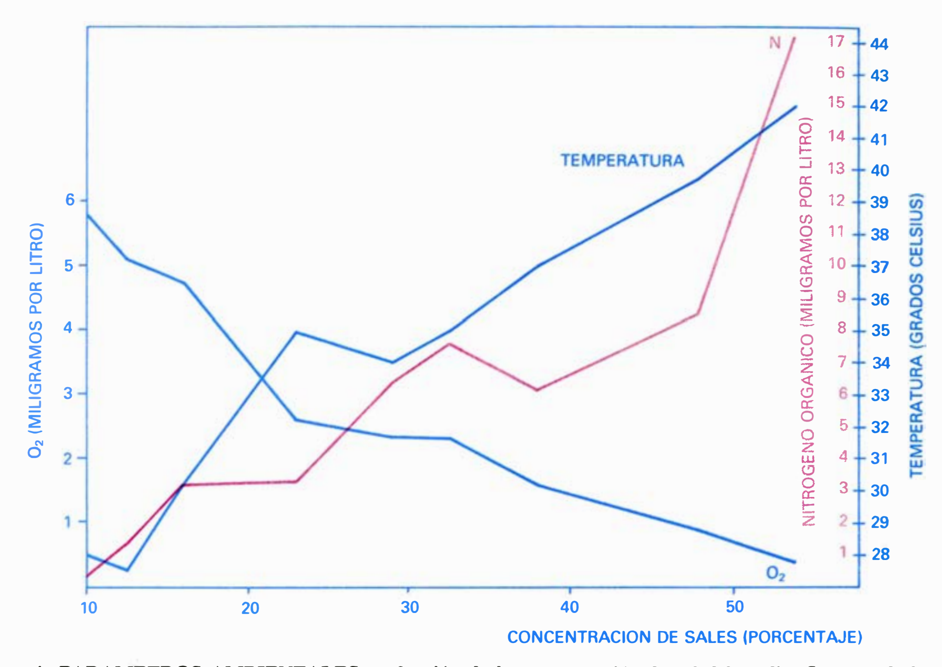
3. SALINA SOLAR de estanque múltiple (Santa Pola, Alicante). En la fotografía se observa un estanque donde se evapora el agua de mar. Estas salinas constituyen un modelo seminatural casi perfecto para el estudio de los organismos halófilos. El sistema permite disponer simultáneamente de una serie de balsas de distinta concentración de sales, desde ligeramente superior al agua de mar hasta un 50 por ciento en sales, y que varía poco en el tiempo.

exterior mayor, la célula perderá agua y se deshidratará hasta igualar la concentración externa, no existiendo ninguna fuerza que evite este proceso sin cerrarle, al mismo tiempo, el paso a las moléculas de agua. Este es un problema que la vida en nuestro planeta no ha resuelto todavía sin cerrar el paso a nutrientes fundamentales para el metabolismo. La solución universal la constituyen los llamados solutos compatibles; con ellos se igualan las concentraciones intracelulares a las exteriores sin perturbar el metabolismo celular. La naturaleza de estos solutos compatibles es muy variable, siendo un ejemplar típico el glicerol acumulado por el alga *Dunaliella* para igualar la concentración exterior de sales en las salinas. En cualquier caso, incluso el glicerol, a altas concentraciones, disminuirá notoriamente la actividad de agua dentro de la célula representando un inevitable enlentecimiento de todas las reacciones metabólicas. Como este es un problema que afectará también a los otros posibles competidores, todo se remite a un enlentecimiento general. Así pues, los hábitats hipersalinos son lo que, en términos cinematográficos, denominaríamos un ecosistema a cámara lenta.

Además del problema osmótico, los microorganismos halófilos deben resolver otros. Por ejemplo, el relacionado con la naturaleza iónica del soluto. Una bacteria halófila situada en un medio que contenga glicerol o glucosa a concentraciones suficientes –isosmóticas con la óptima de cloruro sódico– lisará (se romperá) en cuestión de segundos. Del mismo modo, una levadura osmofílica que puede crecer a concentraciones saturantes de glucosa, no puede desarrollarse con un 10 por ciento de sales marinas en el medio.

La naturaleza de las sales es fundamental para la supervivencia y metabolismo de las bacterias halófilas. Estas pueden ser de dos tipos: extremas y moderadas. Bacterias halófilas extremas son aquellas que requieren concentraciones de sal muy elevadas, por encima del 15 por ciento. Las moderadas requieren concentraciones mucho más bajas, entre un cinco y un 10 por ciento, para vivir óptimamente. El cloruro sódico es irremplazable tanto para las bacterias halófilas moderadas como para las extremas y también para las bacterias marinas, aunque todas ellas pueden mantener sus células vivas (sin crecer) a concentraciones de magnesio –en forma de cloruros o sulfatos– más bajas que las mínimas de cloruro sódico que permiten a la célula mantener su integridad. Todos estos fenómenos participan de un origen común: el mantenimiento de las estructuras de orden superior de las proteínas. Una proteína está constituida por una, o varias, cadenas polipeptídicas, que son secuencias de aminoácidos que se pliegan en el espacio de una forma perfectamente determinada, fundamental para que puedan desempeñar su función, sea esta estructural o enzimática. Los últimos pliegues que terminan de conformar la estructura de las proteínas se deben a dos tipos principales de fuerzas: electrostáticas e hidrofóbicas. Las primeras dependen de ciertos aminoácidos que llevan en su cadena lateral (la parte de la molécula que sobresale de la sucesión de enlaces peptídicos a modo de cola) grupos ionizables de naturaleza ácida ($-COOH$) o básica ($-NH_2$). Estos grupos, cuando están ionizados, poseen carga eléctrica e interaccionan entre sí por atracción o repulsión, conformando la molécula de una manera determinada. Las fuerzas hidrofóbicas se deben también a la diferente naturaleza de los radicales de los aminoácidos, que a veces tienen poca afinidad por el agua (tipo hidrofóbico) y tienden a sepultarse en el interior de la molécula, lo que resulta termodinámicamente más estable.

Ambos tipos de interacciones dependen estrechamente de la fuerza iónica del medio y del tamaño de los iones. Las cargas de los aminoácidos se neutralizan por los iones de carga opuesta; además, la formación de agregados moleculares en el agua cambia en las soluciones de sales y esto modifica sus propiedades como disolvente. El resultado final es que una determinada proteína tiene unos requerimientos especiales en cuanto a los valores de estas dos fuerzas con las que puede mantener su estructura; cuando caen fuera de ciertos intervalos, es inevitable y a veces irreversible la modificación estructural, lo que implica la pérdida de las funciones correspondientes. Otro aspecto importante, aunque hasta el momento más oscuro, es el relativo a la existencia de un entorno iónico adecuado para los fenómenos de transporte activo. Algunas bacterias marinas no pueden transportar moléculas de glucosa hasta el interior de la célula si no existe en el exterior una concentración adecuada (parecida a la del mar) de iones sodio. Cuál sea la razón se desconoce, aunque quizá tenga que ver con el funcionamiento de las permeasas correspondientes.



4. PARAMETROS AMBIENTALES en función de la concentración de sal del medio. Se trata de la temperatura, concentración de oxígeno y nitrógeno orgánico y amoniacal; el último puede considerarse una estimación de la biomasa total existente. Como muestra la gráfica, cuanto mayor es la concentración de sal del medio, mayor es la temperatura y menor la concentración de oxígeno. Esto puede explicarse apelando al menor calor específico del agua y la disminución de la solubilidad del oxígeno al aumentar la concentración de sales. El aumento que se observa en la biomasa total se justifica por la concentración de los nutrientes que permiten una elevada productividad en los estanques muy concentrados. Los datos que se expresan aquí corresponden a muestras obtenidas entre las 13 y 14 horas de un día soleado de agosto.

Los microorganismos halófilos o halotolerantes deberán enfrentarse con los problemas mencionados. En el caso de los halófilos extremos, la adaptación a las concentraciones hipersalinas es tan profunda que cabría preguntarse si se trata de organismos adaptados a altas concentraciones de sal o

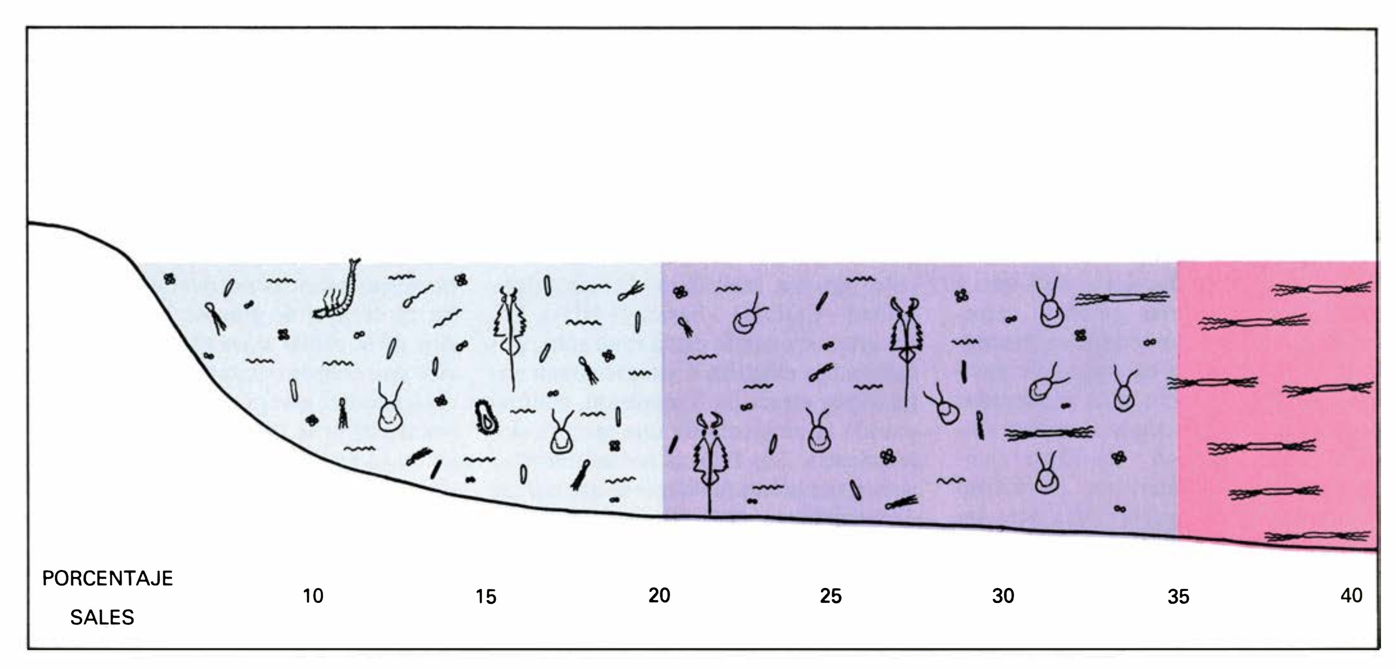
son todos los demás seres vivos los que están àdaptados para vivir sin sales. En eucariotas halófilos, algas y levaduras, la solución adoptada para estos problemas es la de acumular polioles como solutos compatibles. (Los polioles son moléculas orgánicas con varios grupos OH.) En el alga *Dunaliella* la concentración de glicerol en el interior de las células varía directamente con la concentración de sal en el exterior. La actividad de los enzimas –no sólo de *Dunaliella*, sino de eucariotas no halófilos como *Saccharomyces cerevisiae*– se ve muy poco afectada por altas concentraciones de estos solutos, salvo en la aparición de un progresivo enlentecimiento, lo que parece ser una consecuencia inevitable de la pérdida de actividad de agua. El mecanismo por el que se regula la concentración de glicerol intracelular en función de la exterior de sales no se ha investigado todavía, aunque se ha propuesto que podría implicar la selección entre dos vías alternativas del metabolismo del CO_2 fijado fotosintéticamente. *Dunaliella* tiene, además, una elevada capacidad para bombear sodio al exterior y concentrar ligeramente el potasio.

De todos estos hechos parece deducirse que *D. viridis*, que tiene parientes marinos próximos que también acumulan glicerol y, aunque en menor grado, expulsan sodio, no ha sufrido grandes modificaciones para transformarse en un halófilo que crece hasta concentraciones muy elevadas, de hasta un 30 por ciento de sales en el medio. La membrana plasmática del alga podrá aislar eficazmente su medio intracelular, pero algunos enzimas, como las permeasas encargadas de regular la entrada de nutrientes en el interior de la célula, se verán expuestas al medio exterior. Estas proteínas debieron adaptar sus estructuras a una exposición parcial al medio de alta fuerza iónica; probablemente son las responsables de que el alga no crezca por debajo de un ocho por ciento de cloruro sódico en el medio.

Los procariotas halófilos son mucho más variados en su fisiología. Se han dividido entre halófilos moderados y extremos, en función de sus requerimientos para la sal. Se trata de una razón clasificatoria con cierto valor ecológico; y, al menos en lo que respecta a los heterótrofos aerobios, parece que corresponde a microorganismos con hábitats y estrategias de supervivencia distintos. Desde el punto de vista taxonómico, los halófilos moderados representan un conjunto heterogéneo que incluye bacterias gram positivas y gram negativas (los dos grandes troncos en la evolución de las bacterias, que se distinguen por la estructura de su pared celular).

Dado que estos grupos han evolucionado siguiendo caminos independientes, no se pueden hacer generalizaciones respecto a sus mecanismos de adaptación al medio hipersalino. La única característica que comparten, además de su hábitat, es la de precisar concentraciones de sal entre un 5 y un 15 por ciento para funcionar óptimamente. Se decía de ellos que dependían por entero de la sal; sin embargo, aunque esto sea cierto para la mayoría de los que viven en aguas hipersalinas, hemos comprobado que, en suelos hipersalinos, los halófilos moderados presentes, a pesar de tener óptimos alrededor del 10 por ciento, se desarrollan también a concentraciones muy bajas, del orden del 0,5 por ciento. Lo que tiene su explicación: una lluvia abundante en un suelo por hipersalino que sea rebajará la concentración de sales.

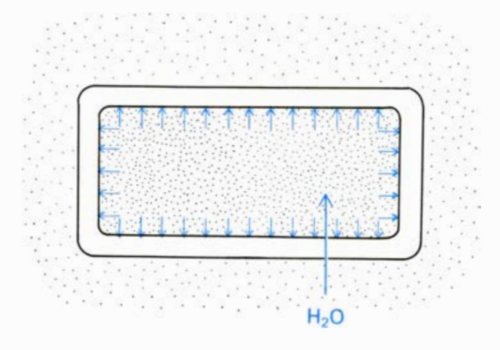
Casi todo lo que se sabe sobre halófilos moderados se refiere a un pequeño vibrión de la especie *Vibrio costicola*, una bacteria gram negativa del mismo grupo que el agente productor del cólera y que está muy bien representada en ambientes marinos. Esta bacteria crece bien hasta con un 15-20 por ciento de sales en el medio. (Hay halófilos moderados que soportan concentraciones bastante más elevadas, aunque ninguno suele superar el 30 por ciento.) *Vibrio costicola* necesita el sodio para crecer y en agua destilada lisa masivamente en unos segundos. En agua de mar prolifera muy despacio, pero sobrevive sin mayores problemas. Se sabe muy poco del soluto compatible de los halófilos moderados; acumulan potasio en el interior y expulsan sodio, fenómeno que parece ser universal en todos los seres vivos. Sin embargo, la cantidad de potasio intracelular no es suficiente para equilibrar la concentración externa.

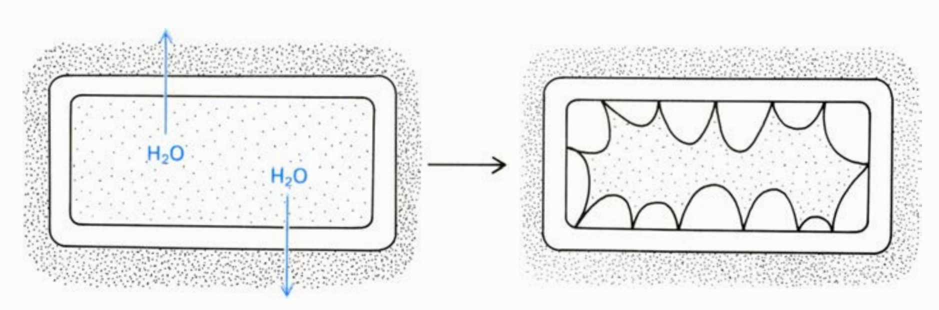


5. DIVERSIDAD de las poblaciones en aguas hipersalinas con concentraciones diferentes de sal. Entre 10 y 20 por ciento de sales, abundan las bacterias halófilas moderadas con morfologías muy distintas: bacilos, cocos y espirilos, protozoos, algas y pequeños crustáceos. A partir del 20 por ciento la diversidad cae: disminuyen los espirilos, permanecen los cocos y bacilos halófilos moderados, *Artemia salina* y *Dunaliella*. Al final de este intervalo aparecen halófilos extremos: cocos y bacilos. A partir del 30 por ciento sólo restan halófilos extremos y *Dunaliella*. A partir del 35 por ciento, el alga desaparece.

Además, la concentración de potasio no varía de una forma directamente proporcional con la de sales del exterior. ¿Cuál es, pues, el soluto compatible de los halófilos moderados? Es posible que cuando se encuentran en medios con pocas sales, en las primeras etapas de concentración del agua de mar, la acumulación de potasio intracelular junto con el sodio que toleran intracelularmente –la mayor concentración detectada en *Vibrio costicola* ha sido del orden de 1 molar de cloruro sódico equivalente a un seis por ciento– sean suficientes. A concentraciones más altas deben recurrir a algún otro soluto compatible. Se han propuesto aminoácidos neutros, como la prolina, y otras pequeñas moléculas orgánicas, pero en ningún caso ha sido demostrado. Por lo que sabemos, los diferentes halófilos moderados pueden haber desarrollado distintos solutos compatibles, aunque en general su entorno iónico intracelular nunca es muy extremo. (Del estudio de sus enzimas se desprende que nunca funcionan a altas concentraciones de sal.)

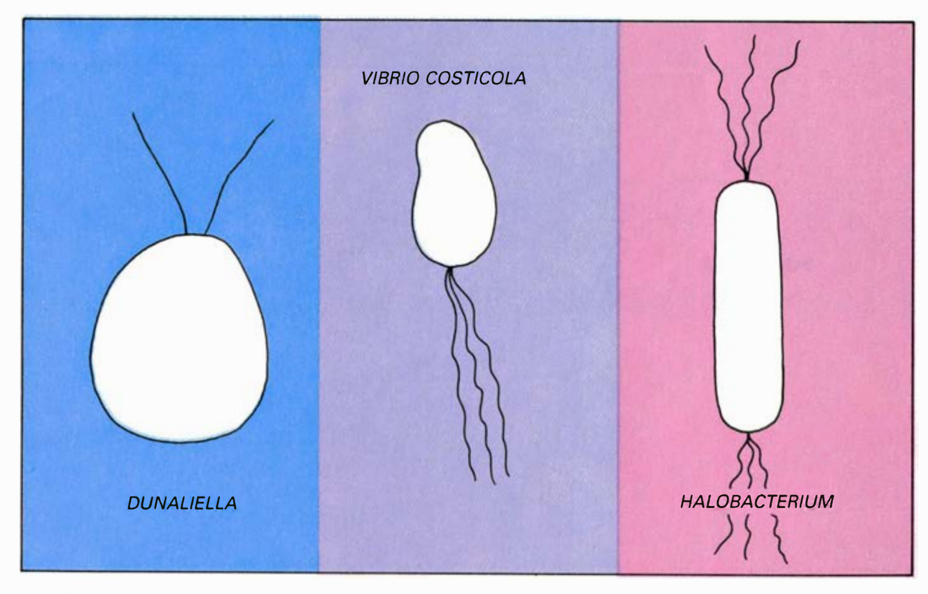
Además de los componentes de la familia Halobacteriáceas, algunas bacterias fotosintéticas del género *Ectothiorhodospira* se consideran también halófilos extremos, puesto que crecen dentro de un intervalo de concentraciones de sal parecido. Estos organismos realizan una fotosíntesis anaerobia utilizando ácido sulfhídrico (SH_2) como donador de electrones; dependiendo por tanto del suministro de éste, que captan de los lodos anaeróbicos. Sabemos muy poco de los hábitats anaerobios hipersalinos. En general, como ocurre con los halófilos moderados, no son muy distintos de sus parientes taxonómicos no halófilos.



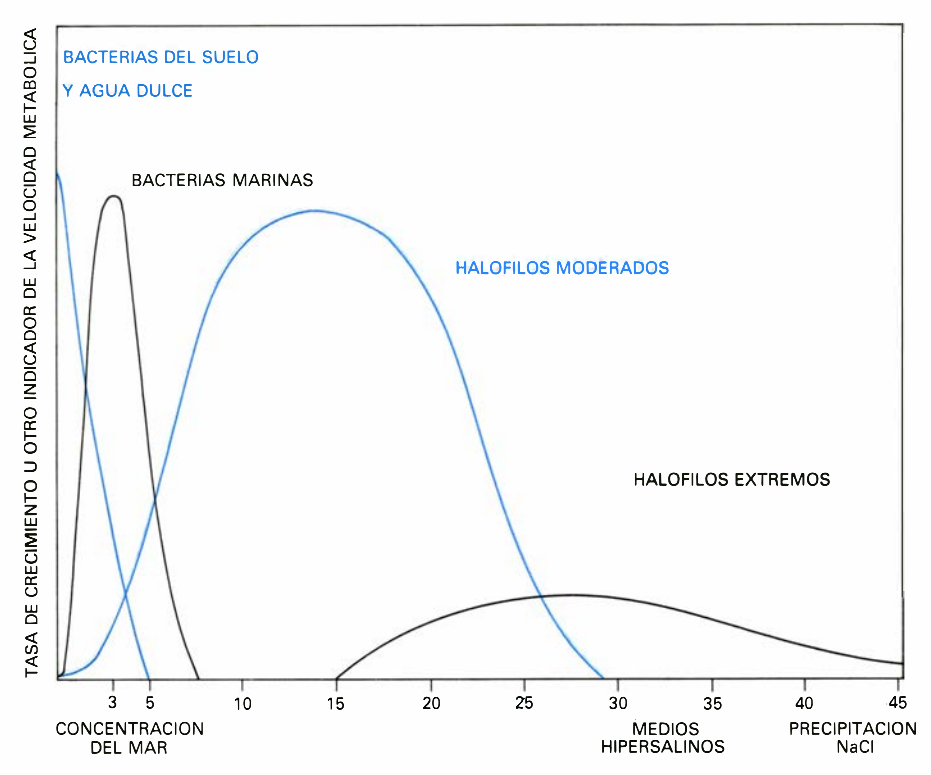


6. EFECTO DE LAS CONCENTRACIONES EXTERNAS de solutos sobre las células. Cuando la concentración externa es menor, el agua penetra en el interior de la célula y crea una presión positiva que es absorbida por la pared celular rígida. Cuando la concentración externa es mayor, el agua sale de la célula produciendo la llamada plasmolisis: el citoplasma comienza a retraerse y la célula acaba por morir.

Las halobacterias son únicas en muchos aspectos. Utilizan un soluto compatible iónico –el ion potasio– y modifican su concentración adaptándola a la exterior. En *Halobacterium* se han detectado concentraciones intracelulares de potasio del orden de 5 M, lo que representa un 37 por ciento de cloruro potásico; y esto ocurre aunque la concentración de potasio en el exterior sea cientos de veces más pequeña. Hay, por contra, muy poco sodio intracelular: alrededor de 1 M. Para neutralizar estos iones positivos, existen aniones cloruro; sin embargo, a duras penas logran neutralizar la mitad de las cargas de los iones positivos. Los demás aniones los constituyen las propias proteínas. Las enzimas y las proteínas estructurales de los halófilos extremos son ácidas; es decir, contienen un exceso de aminoácidos ácidos que portan un grupo $-COOH$ extra. Al ionizarse, este grupo adquiere carga negativa y convierte la proteína halófila en un polianión. Las curvas de actividad de los enzimas halófilos estudiados demuestran que requieren cationes como el sodio o el potasio para efectuar su función catalítica, preferentemente, como era de esperar, potasio. Por otra parte, la exposición de una proteína halófila a condiciones de muy baja fuerza iónica conlleva muchas veces su desnaturalización

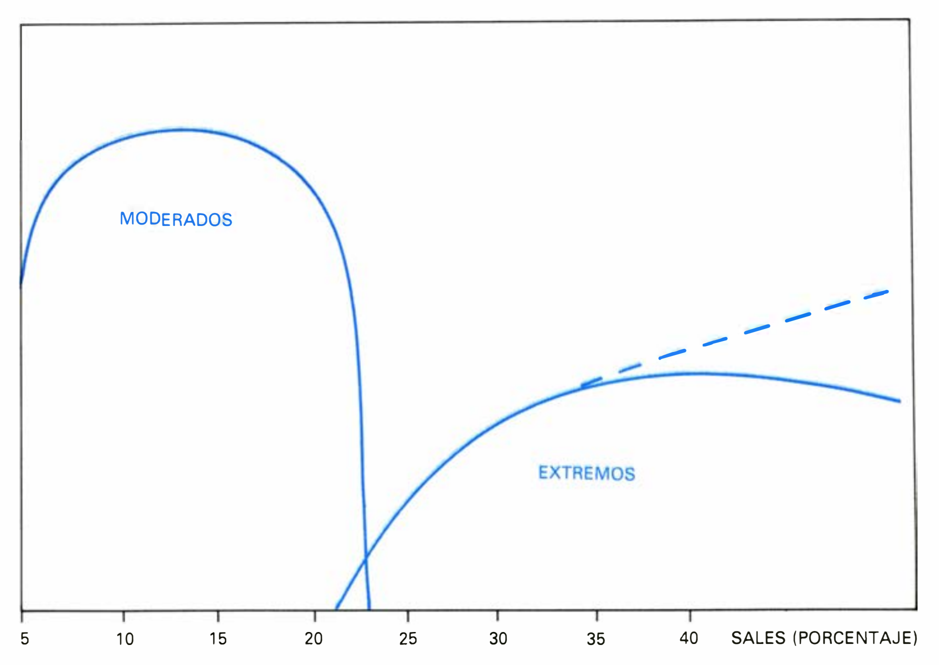


7. SOLUTOS COMPATIBLES usados por diversos organismos halófilos para compensar la elevada concentración de sales en el exterior. Estos solutos no afectan al metabolismo de las células y permiten igualar la concentración externa. En el alga *Dunaliella* es el glicerol (*izquierda*). En el halófilo moderado *Vibrio costicola* es, parcialmente, el potasio y el sodio (*centro*). Sin embargo, cuando las concentraciones extracelulares son muy altas, estos últimos no pueden ser tolerados a niveles tan elevados. Entra en juego entonces otro soluto compatible cuya naturaleza exacta se desconoce; podría tratarse de un compuesto orgánico neutro como el aminoácido prolina. En los halófilos extremos, el soluto compatible es el potasio, que se concentra intracelularmente cientos de veces (*derecha*). El sodio se tolera a concentraciones altas. Los iones positivos son neutralizados por cloruro y por las cargas negativas de las proteínas.



8. TIPOS ADAPTATIVOS con respecto a la concentración de sales marinas que existen entre las bacterias heterotróficas. Se ponen de manifiesto por la velocidad de crecimiento en medios de cultivo con distintas concentraciones de estas sales. Los espectros que se indican pueden considerarse típicos y no corresponden a ningún organismo concreto. Los halófilos moderados y los extremos están especializados para desarrollarse en la vida en ambientes hipersalinos (de salinidad superior a la del mar).

irreversible. ¿Por qué poseen una carga resultante negativa, es decir, por qué son proteínas ácidas? Aunque no tenemos una respuesta tajante, se sabe que las proteínas ácidas son más aptas para funcionar en el medio intracelular de los halófilos extremos. Sorprende la elasticidad que manifiestan los enzimas halofílicos respecto a la concentración de iones del medio, pudiendo funcionar algunos entre 1 y 5 M de potasio. Deben poseer una estructura muy robusta que les permita mantenerse en condiciones tan diferentes.



9. DISTRIBUCION NATURAL DE HALOFILOS MODERADOS Y EXTREMOS en la salina de Santa Pola. Ambos grupos tienen sus nichos bien diferenciados; sólo coexisten en un estrecho intervalo, en estanques que tienen entre un 20 y un 30 por ciento de sales. Aunque esto coincide aproximadamente con los espectros de crecimiento que se muestran en la figura anterior, la competencia entre ambos grupos está afectada por otros factores, tales como la temperatura o la concentración de nutrientes.

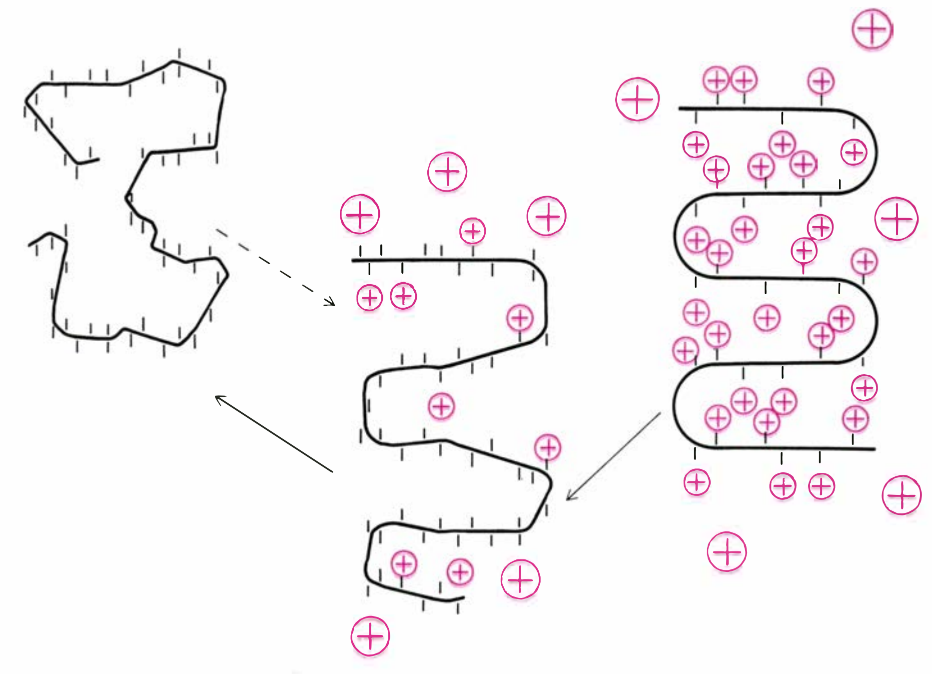
Otro aspecto de la adaptación de las halobacterias a su medio tiene que ver con la radiación luminosa, muy intensa en su medio natural. El color rojo vivo de los halófilos extremos se debe a que poseen carotenoides de 50 átomos de carbono denominados bacteriorruberinas. Estos pigmentos no son muy distintos de los que poseen otros organismos expuestos a fuertes intensidades luminosas; actúan protegiendo a la célula del llamado efecto fotodinámico: muchas moléculas orgánicas pigmentadas, colorantes y componentes normales de las células –en particular los que llevan anillos porfirínicos como citocromos o clorofilas–, cuando están expuestos a fuertes intensidades luminosas, actúan sobre el oxígeno molecular y lo transforman en radicales libres de oxígeno, principalmente en el denominado oxígeno singlete, que presenta dos electrones con espines opuestos en el mismo orbital. El oxígeno singlete, altamente tóxico, es el responsable de la acción germicida de la luz visible. En este sentido, los carotenoides actúan de "extintores", oxidándose y absorbiendo la energía de este radical que pasa a inofensivo oxígeno molecular.

El mecanismo de fotorreactivación que interviene en la reparación del ADN dañado por la radiación ultravioleta también es muy eficaz y parece estar favorecido por los carotenoides. Las halobacterias están bien protegidas de la radiación luminosa; más aún, la utilizan en beneficio propio. La primera prueba en este sentido llegó al comprobar que las vesículas obtenidas por reasociación espontánea de las membranas de células lisadas, cuando se las iluminaba, originaban un descenso del *p*H del medio detectable. Posteriormente se demostró que este curioso fenómeno debíase a una cromoproteína, la bacteriorrodopsina, que tiene el mismo cromóforo que la rodopsina del ojo de los vertebrados: el retinal. Esta proteína la sintetizan en grandes cantidades las células de *Halobacterium halobium* cuando el microorganismo se encuentra en condiciones anaeróbicas. En esta circunstancia llega a constituir buena parte de la superficie total de la célula formando la llamada membrana purpúrea. La bacteriorrodopsina absorbe la energía luminosa y la utiliza para bombear protones del interior de la célula al exterior de la misma; de ahí, el descenso del *p*H observado. Este gradiente de protones puede ser aprovechado para la síntesis de ATP.

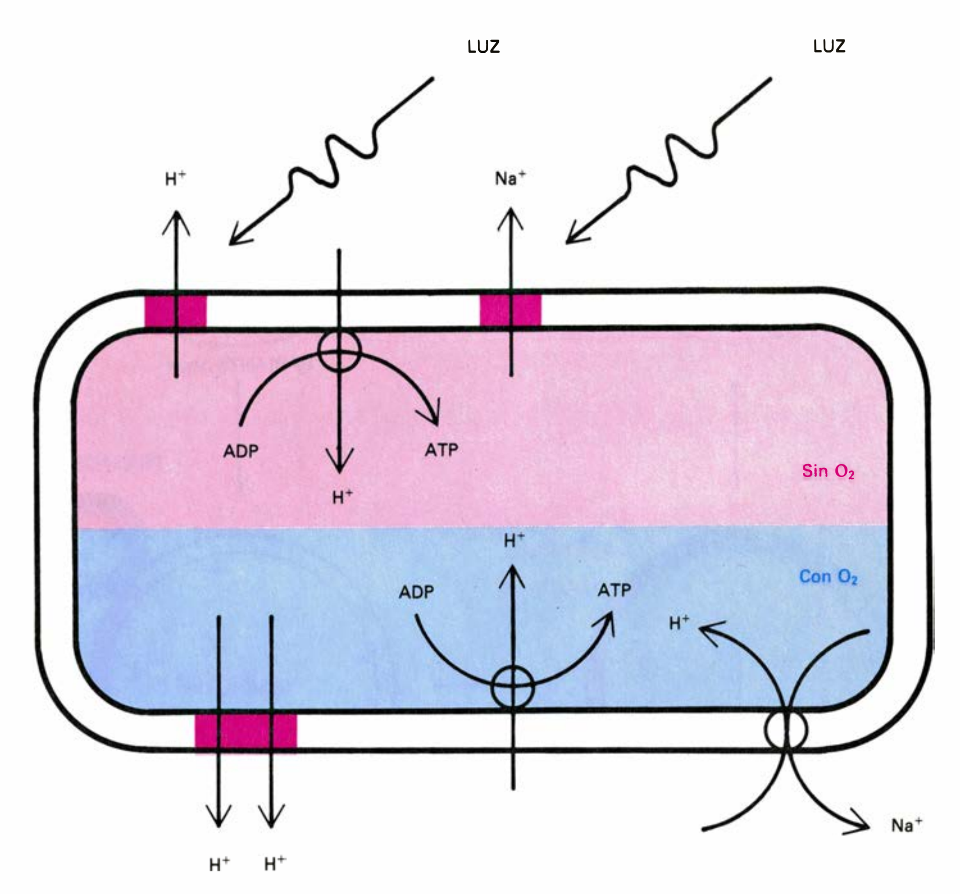
Recientemente se ha descubierto que la bacteriorrodopsina no es el único mecanismo para bombear iones que

poseen las halobacterias. La halorrodopsina, otra proteína con el mismo cromóforo, exporta iones de sodio en presencia de luz, ayudando a mantener el desnivel de concentraciones de sodio entre el interior y el exterior. Hay algo paradójico en torno a la bacteriorrodopsina; al parecer, sólo se sintetiza en condiciones de anaerobiosis que se dirían inhibidoras del desarrollo de *H. halobium*. Se desconoce por qué razón no se sintetiza bacteriorrodopsina en condiciones aerobias, que son las que permiten el crecimiento de las bacterias en el laboratorio. Quizá la clave se encuentre en la misma artificiosidad de las condiciones de laboratorio. En medios naturales, no existen ni las condiciones altamente aerobias ni las elevadas velocidades de crecimiento que se dan in vitro. Ambas pueden considerarse muy poco naturales.

Los procariotas poseen siempre un solo cromosoma, una molécula circular de ADN, donde se encierra toda la información fundamental para la vida de las células. Frecuentemente poseen otras moléculas de ADN más pequeñas: los plásmidos, donde encontramos genes de valor adaptativo, como los que determinan la resistencia a antibióticos o la utilización de sustratos raros. Estos minicromosomas suelen denominarse ADN satélite porque, al centrifugar la mezcla de ADN en un gradiente de concentraciones de cloruro de cesio a gran velocidad, aparecen en una banda distinta del ADN cromosómico. Las halobacterias poseen la proporción más alta de ADN satélite detectada en ningún procariota, próxima en algunos casos al 40 por ciento del total del ADN celular. Esto supone, si el tamaño de los plásmidos de las halobacterias es parecido al de los otros procariotas –y existen datos que indican que realmente es así– entre ocho y 20 plásmidos por célula. ¿Qué significa tan abultada cantidad de ADN extracromosómico? F. Pfeifer y colaboradores, de la Universidad de Würzburg, han estudiado la relación entre algunos plásmidos caracterizados en *Halobacterium* y la variabilidad espontánea de caracteres fenotípicos, concretamente la formación de vacuolas de gas, que es una característica de algunas cepas de *H. halobium*. Se sabía que esta característica se hallaba ligada a un plásmido denominado pHH1. Las cepas que no presentan este plásmido nunca forman vacuolas de gas. En cualquier caso, siempre existe en las poblaciones de esta bacteria una proporción alta –alrededor de un 1 por ciento– de células sin vacuolas de gas. Usando técnicas de



10. FENOMENO DEL DESALADO (o "salting out") de las proteínas halófilas, propuesto por Janos Lanyi, del Ames Research Center en California. Las proteínas de los halófilos extremos contienen un sobreexceso de aminoácidos ácidos que, al ionizarse, se cargan negativamente. La conformación de la proteína es estable mientras existan cationes que neutralicen estas cargas. Los cationes han de ser de cierto tamaño para poder penetrar en los recovecos de la proteína. Al ser expuestas a concentraciones bajas de sal, algunas cargas quedan sin neutralizar, la estructura se convierte en inestable y se desnaturaliza, a veces irreversiblemente, la proteína. Las fuerzas hidrofóbicas que estabilizan la molécula por la tendencia de ciertos grupos a alejarse del agua también se modifican al variar la concentración de sales.



11. DUALIDAD DE LOS MECANISMOS de producción de energía (ATP) y de expulsión de sodio en *Halobacterium*. Sin oxígeno, la energía luminosa puede ser utilizada por la bacteriorrodopsina y halorrodopsina de la membrana purpúrea para la expulsión activa de protones e iones sodio. Un enzima ATPasa aprovecha luego el gradiente de protones para producir ATP. En condiciones aeróbicas, la cadena respiratoria crea el gradiente de protones que puede utilizarse también para sacar sodio por un transportador de sentido inverso ("antiporter"): transportador de sodio dependiente de la entrada de protones.

análisis del ADN plasmídico con restrictasas, Pfeifer ha comprobado que esta variabilidad se debe a la inserción de transposones en el plásmido; lo más sorprendente de sus hallazgos es la elevada frecuencia con que aparecen inserciones en los plásmidos estudiados. Ello revela un material genético sometido a un cambio continuo, donde muchos genes y grupos de genes se activan e inactivan.

Se está popularizando la expresión "genoma elástico" para referirse al genoma de las halobacterias, en razón de su extremada plasticidad. Todas las bacterias poseen pequeñas bibliotecas de genes que no utilizan normalmente, pero que activan cuando las condiciones del medio lo requieren. En las halobacterias, esta biblioteca puede ser sumamente rica. Con el mismo tipo de técnicas, C. Sapienza y W. F. Doolittle, de la Universidad de Dalhousie, acaban de detectar la presencia de varias familias de secuencias repetitivas en los genomas de dos halobacterias: *Halobacterium halobium* y *H. volcanii*. Existen numerosas familias –50 por lo menos– que contienen de dos a 20 copias de cada secuencia. Semejante fenómeno se ha descubierto también en algunas plantas y animales, pero nunca en procariotas. De donde se deduce que la organización genética de estos microorganismos difiere de la que caracteriza a los procariotas típicos.

¿De dónde les viene a las halobacterias su singularidad? ¿Resulta acaso de un largo proceso adaptativo? Sí, al menos en parte. Pero los halófilos moderados también están adaptados a los medios hipersalinos. Compiten con los extremos de un modo eficaz. En muy buena parte, la singularidad de las halobacterias es consecuencia de su filogenia. Pertenecen a un tronco que se desgajó de los demás procariotas hace miles de millones de años, cuando caracteres tan básicos para los procariotas como la naturaleza de la pared celular estaban todavía en proceso de formación. Es probable que las halobacterias lleven más de 3000 millones de años evolucionando por su cuenta en su casi exclusivo hábitat.

Sobre sus relaciones con otras archibacterias, y basándose en las secuencias de los ARN 16 S, ribosómicos, se ha sugerido que representan una rama colateral de los metanógenos, si bien las diferencias entre ambos grupos son enormes. (Los metágenos son archibacterias extremadamente anaeróbicas que producen metano.) Los estudiosos del origen de la vida han sugerido muchas veces que los medios en que surgieron los primeros organismos podrían ser estanques o lagunas junto al mar, donde los compuestos orgánicos se concentrarían por evaporación. Si entonces el mar era ya salado –y probablemente lo era en mayor o menor grado– las sales también se concentrarían. Estos medios serían hipersalinos, anaeróbicos y cálidos. ¿Han sobrevivido las archibacterias en medios que recuerdan los remotos orígenes de la vida? Hay un punto interesante que relaciona las archibacterias, y más concretamente las halobacterias, con el origen de la vida: sus mecanismos de obtención de energía. La bacteriorrodopsina puede considerarse como el método más simple de transformación de energía solar en ATP que existe, y puede haberse originado evolutivamente en un tiempo muy corto. Se creía antes que los primeros seres vivos eran microorganismos anaerobios que oxidarían la materia orgánica por medio de las vías fermentativas (relativamente ineficaces y despilfarradoras), en las que unos compuestos orgánicos se utilizan como aceptores electrónicos (por ejemplo el ácido pirívico) en la oxidación de otros (verbigracia la glucosa). En un medio sin productores primarios que repongan la materia orgánica utilizada por los heterótrofos, este es un comportamiento ecológicamente suicida, porque implicaría un continuo empobrecimiento del hábitat. Cabe preguntarse cuánto tiempo podría haber durado un ecosistema de este tipo, incluso admitiendo la acumulación de cantidades muy grandes de materia orgánica de origen abiótico. Es dudoso que estos organismos hubieran tenido tiempo de desarrollar mecanismos de obtención de energía del tipo de la fotosíntesis.

Posiblemente un mecanismo como el de la bacteriorrodopsina hubiera resuelto temporalmente el problema; las reservas de materia orgánica abiótica se aprovecharían sólo como fuente de carbono, consumiéndose, por tanto, mucho más despacio, siendo la luz la fuente de energía. También los mecanismos de obtención de energía de los metanógenos serían adecuados en una situación de este tipo, ya que pueden oxidar el hidrógeno molecular utilizando CO_2 como aceptor de electrones, compuestos ambos que debían abundar en la atmósfera primitiva. Estos mecanismos sencillos de obtención de energía quizá sean los últimos vestigios de las fuentes de energía de los primeros ecosistemas.

FRECUENCIA: 10^{-4} · DELECCION · GENES *Vac* · PHH1 · VARIANTE *Vac*$^{+}$ · TRASPOSON · INSERCION · GENES *Vac* INACTIVOS · VARIANTE *Vac*$^{-}$ · FRECUENCIA: 10^{-2}

12. MECANISMO DE ALTA VARIABILIDAD FENOTIPICA en *Halobacterium halobium* mediado por transposones, propuesto por F. Pfeifer y colaboradores. Los genes *Vac* responsables de la formación de vacuolas de gas que regulan la flotabilidad del microorganismo se encuentran en el plásmido pHH1. La inserción de un transposón en el plásmido conlleva la inactivación de estos genes, originando colonias *Vac*. Esto ocurre en una de cada cien células. A su vez, una de cada 10.000 con el transposón insertado lo pierde por deleción y origina una colonia *Vac*$^{+}$. Estos fenómenos no son raros en las halobacterias.

Temas metamágicos

Recursivando retozonamente en Lisp, lenguaje de la inteligencia artificial

Douglas R. Hofstadter

Ya que el mes pasado terminé aludiendo oportunamente al hogareño cuerpoespín glazunkiano, bueno puede ser comenzar ahora añadiendo unas palabras acerca de tan notable como desconocida bestezuela. Como sin duda recuerda el lector, las púas de todos los cuerpoespines (a excepción de los de tamaño mínimo) son cuerpoespines más pequeños. Los cuerpoespines mínimos carecen de púas; tienen, en cambio, hocico, y por cierto que valioso, pues los glazunkianos han fundado sobre tal objeto todo su sistema monetario. Fijémonos en el valor de los cuerpoespines de dimensión tres, de la variedad que habita en Glazunkia Exterior. Cada cuerpoespín de esta variedad cuenta con nueve púas (en contraste con sus primos de Glazunkia Interior, que tienen siete); por tanto, de su cuerpo emanan nueve cuerpoespines de dimensión dos, cada uno de los cuales se ramifica en otros nueve de tamaño uno. Por fin, cada uno de éstos es punto de partida de otros nueve cuerpoespines de dimensión cero, los cuales solamente tienen hocico. Contando todo, resultan $9 \times 9 \times 9 \times 1$ hocicos, lo que significa que un cuerpoespín de tamaño tres de Glazunkia Exterior tiene una capacidad adquisitiva de 729 hocicos. Si, por otra parte, nos encontrásemos en Glazunkia Interior, y comenzásemos con un cuerpoespín de tamaño 4, su capacidad de compra sería $7 \times 7 \times 7 \times 7 \times 1$, o sea, 2041 hocicos.

A ver si logramos atinar con una receta general que permita calcular el poder adquisitivo de un cuerpoespín cualquiera. Podría servirnos algo así:

El poder adquisitivo de un cuerpoespín de dimensión dada, que tenga determinado recuento de púas, es:
- si la dimensión es 0, entonces el poder adquisitivo es 1;
- en cualquier otro caso, se calcula la potencia adquisitiva de un cuerpoespín de idéntico recuento y tamaño inmediatamente inferior, y se multiplica por el número de púas.

Podemos abreviar la receta ayudándonos de cierto simbolismo. Ante todo, denotemos "p" al número de púas que emanan de cada cuerpoespín, y sea "d" la dimensión, o tamaño, del animalejo. En seguida, expresemos mediante la partícula "cond" el "si" condicional, y sea "t" (de "true", verdadero) la abreviatura de "en cualquier otro caso". Nos valdremos además de una especie de notación algebraica condensada, en la cual los nombres de las operaciones se escriben a la izquierda de sus operandos, entre paréntesis:

```
(potencia-adquisitiva p d) es:
  cond (eq d 0) 1;
    t (producto
        p
        (potencia-adquisitiva
          p
          (inmediata-inferior d)))
```

Tenemos aquí una transcripción exacta de la receta antes dada a un lenguaje algo más simbólico. Podemos hacerlo todavía más simbólico y compacto gracias a un par de nuevos convenios. Encerremos entre paréntesis cada uno de los dos casos (el caso en que d sea igual a 0, y el "cualquier otro" caso). En general, se usan paréntesis cada vez que se desea delimitar perfectamente una unidad lógica. Por fin, indicaremos mediante las palabras "def" y "lambda" que estamos definiendo una función general, llamada "potencia-adquisitiva" que depende de dos variables, el número de púas, p, y el tamaño o dimensión, d. Obtenemos ahora:

```
(def potencia-adquisitiva
 (lambda (p d)
 (cond ((eq d 0) 1)
   (t
    (producto p
       (capacidad-adquisitiva p
          (inmediata-inferior d)))))))
```

Ya he mencionado anteriormente que la capacidad adquisitiva de un cuerpoespín de tamaño tres y nueve púas por tamaño es de 729 hocicos. Podemos expresarlo ahora diciendo que (potencia-adquisitiva 9 3) es igual a 729. Análogamente, (potencia-adquisitiva 7 4) será igual a 2041.

Retornemos al Lisp. Hacia el término de la sección del mes pasado había yo propuesto el problema de redactar en Lisp una función que englobara a toda una familia de funciones llamadas "cuadrado", "cubo", "cuarta potencia", "quinta potencia", etcétera. Pedía yo a los lectores que construyesen una función general, llamada "potencia", dependiente de dos variables, tal que "(potencia 9 3)" diese 729, "(potencia 7 4) diese 2041, y así sucesivamente. Había yo presentado una "torre de potencias", es decir, una torre infinitamente alta de definiciones individuales en Lisp, una para cada potencia, que conectaba cada definición con la definición de la potencia precedente. Una de las plantas típicas de esta torre podría ser:

```
(def 102-ésima-potencia
 (lambda (p)
  (producto p (101-ésima-potencia p))))
```

Evidentemente, la 101-ésima potencia estaría definida con auxilio de la 100-ésima potencia, y así sucesivamente, creando de esta forma una regresión bastante larga, aunque finita, hasta alcanzar el caso "embrionario", esto es, el caso más sencillo. Incidentalmente, este caso de sencillez extrema, más que el "cuadrado" o la "primera potencia", es éste:

```
(def 0-ésima-potencia (lambda (p) 1))
```

Ya dije el mes pasado a los lectores que disponían entonces de toda la información necesaria para ensamblar la definición requerida. Todo cuanto hay que observar es que cada nivel de la torre descansa sobre el nivel "inmediatamente inferior", a excepción del "nivel de base", que se sostiene a sí mismo. Al decir "inmediatamente inferior" quiero decir lo siguiente:

```
(def inmediato-inferior
  (lambda (d) (diferencia d 1)))
```

Así pues, "(inmediato-inferior 102)" produce un retorno igual a 101. En realidad, la función "inmediato-inferior" dispone en Lisp de nombre propio, a

saber, "sub 1" (substraer 1); y hay también nombre especial para su función recíproca, consistente en añadir 1:"add 1". Ensamblando todas las observaciones anteriores llegamos a la siguiente definición universal:

```
(def potencia
  (lambda (p d)
    (cond ((eq d 0) 1)
      (t
        (producto p
          (potencia p
            (inmediata-inferior d)))))))
```

Esta es la solución al rompecabezas que propuse. Humm... Es curioso. Tengo una extraña sensación de *déjà vu*. Me pregunto por qué.

La definición aquí presentada es una definición *recursiva*, por la razón de que se usa el definiendum en el seno del definiens. Que es una forma cultista de decir que lo definido entra en la definición, que estoy definiendo algo refiriéndolo a sí mismo; y esto parece ser desmañado cuando no manifiestamente circular. Para ver si el geniecillo de Lisp mira con desconfianza semejantes artificios, pidámosle que calcule (potencia 9 3):

```
— > (potencia 9 3)
729
— >
```

¡Esta sí que es buena! ¡No hay protesta! ¿Cómo es que el genio del Lisp puede tragarse semejantes absurdos?

La mejor explicación que puedo dar es que realmente no hay en la definición círculo vicioso alguno. Aunque es cierto que la definición de "potencia" contiene la palabra "potencia", las dos ocasiones en que aparece esta palabra se hace alusión a circunstancias diferentes. En una píldora, (potencia p d) está definida por intermedio de un caso más sencillo, a saber, (potencia p (inferior-inmediata d)). Así pues, estoy definiendo la cuadragesimocuarta potencia por intermedio de la cuadragesimotercera; a ésta, por intermedio de la inmediatamente inferior, y así sucesivamente hasta llegar a la "planta baja", como he venido llamándola: la potencia 0-ésima, que está definida explícita y no recursivamente. Basta entonces decirle al genio que tal valor es 1, y así lo hemos hecho. Por tanto, si se observa cuidadosamente, esta definición recursiva es tan poco circular como pueda serlo la "torre de potencias" –y ¡difícilmente puede algo ser más directo que una línea recta! En efecto, esta compacta definición no es sino una forma de condensar la torre completa de potencias en una expresión finita. Lejos de ser circular, constituye, por el contrario, un cómodo resumen de una infinidad de definiciones diferentes, pertenecientes todas a una misma familia.

En caso de que el lector guarde todavía cierta reserva ante esta especie de prestidigitación, tal vez conviniera dejarle ver lo que haría el Lisp-genio si se le pidiese una "traza" de la función y luego se le rogara nuevamente evaluar (potencia 9 3). Consultemos, para ello, la figura 1.

En las líneas marcadas "ENTERING" el Lisp-genio va imprimiendo los valores de los dos argumentos; y en las líneas rotuladas "EXITING" escribe el valor que ha calculado y que nos retorna. Como es obvio, por cada línea ENTERING hay una línea EXITING, ambas verticalmente alineadas, es decir, sangradas un mismo número de espacios desde el margen.

Podemos ver que para calcular (potencia 9 3) el genio ha de calcular primero (potencia 9 2). Este valor, sin embargo, no es un dato; requiere a su vez conocer el valor de (potencia 9 1), el cual, por su parte, exige conocer (potencia 9 0). ¡Ah, éste sí que nos lo han dado! Es, sencillamente, 1. Y ahora podemos ir "ascendiendo", recordando que para ir obteniendo cada respuesta a partir de la más "profunda" tenemos que ir multiplicando por 9. Así obtenemos primero 9, y luego 81 y después 729, con lo cual concluimos el cálculo.

He dicho "concluimos", pero desde luego no somos nosotros, sino el Lisp-genio, quien debe llevar la cuenta de estas cosas. El Lisp-genio ha de ser capaz de suspender un cálculo y dedicarse a trabajar en otro cuya solución sea necesaria para el primero. Y el segundo cálculo puede muy bien requerir la solución de un tercero, y por tanto, de poder quedar a la espera, y lo mismo este tercero, etcétera, y así iterativamente. Empero, habrá eventualmente de llegarse a un caso donde la profundización se detenga –es decir, un proceso que llegue a completarse efectivamente, retornando un valor– y ello permitirá que los otros procesos aplicados y pendientes puedan finalmente ir retornando valores; es algo así como los aviones a la espera, "en circuito" sobre los aeropuertos grandes, que a veces han de permanecer durante horas para poder aterrizar, dejando cada aterrizaje la pista libre para el siguiente de la cola.

Ordinariamente el Lisp-genio no va dejando traza de lo que va pensando, a menos que así se lo solicitemos. Pero tanto si pedimos verlo como si no, se va desarrollando entre bastidores un proceso de esta clase cada vez que se evalúa una función. Uno de los aspectos más gratos de Lisp es que pueda manejar estas largas definiciones recursivas sin aturdirse.

No llega mi ingenuidad al extremo de suponer que con esto habrán los lectores captado por completo la noción de recurrencia, encontrándose ya listos para escribir con soltura largos programas recursivos. Más aún, la recurrencia suele ser un procedimiento harto sutil de definir funciones, y no es excepcional que hasta los expertos tropiecen con dificultades a la hora de hallar el significado de una función recursiva complicada. He pensado, por consiguiente, que podría ir bien a los lectores un poco de práctica en el trabajo por recurrencia.

Permítaseme presentar un sencillo ejemplo basado en esta charada: ¿Cómo se construye un montón de 13 piedras? Una respuesta "tonta" podría ser: poniendo una piedra encima de otras 12. Supongamos que nos convenga crear una función, definida en Lisp,

```
—> (potencia 9 3)
    ENTERING potencia (p = 9, d = 3)
      ENTERING potencia (p = 9, d = 2)
        ENTERING potencia (p = 9, d = 1)
          ENTERING potencia (p = 9, d = 0)
          EXITING    potencia    (valor: 1)
        EXITING    potencia    (valor: 9)
      EXITING    potencia    (valor: 81)
    EXITING    potencia    (valor: 729)
729
—>
```

1. El genio del Lisp evalúa (potencia 9 3)

que nos proporcione, no un montón de 13 piedras, sino una lista consistente en 13 copias del átomo "piedra" o, generalizando, *n* copias de ese átomo. Podemos fundar nuestra solución en la respuesta, correcta, aunque aparentemente tonta, de la charada. La noción general consiste en construir la respuesta correspondiente al caso *n* a partir de la construida para el antecesor de *n*. ¿Construirla cómo? Pues usando la función cons, de construcción de listas. ¿Cuál es el caso embrionario? Es decir, ¿para qué valor de *n* carece por completo de dificultad la pregunta? Eso es fácil: cuando *n* sea igual a 0 nuestra lista deberá estar vacía, lo que significa que la respuesta es nil. Podemos ahora organizar nuestras observaciones así:

```
(def montón-de-piedras
 (lambda (n)
  (cond ((eq n 0) nil)
    (t
     (cons 'piedra
      (montón-de-piedras
         (inmediato-inferior n)))))))
```

Fijémonos ahora en cómo el genio va acumulando un montoncito de piedras muy pequeño (con "traza" en funcionamiento, para que sea más divertido). A tal fin convendrá que el lector consulte la figura 2.

Tenemos aquí ilustrado lo que hemos llamado "consar una lista". Probemos ahora con otro ejemplo. Se trata de uno de esos chistes gastados del Lisp y, más en general, de la recurrencia. Eche primero un vistazo a la definición, y vea si puede descubrir lo que se supone que hace; siga después leyendo para saber si acertó.

```
(def sorpresa!
 (lambda (n)
  (cond ((eq n 0) 1)
   (t
    (producto n (sorpresa! (sub1 n)))))))
```

Recordemos que "sub1" significa lo mismo que "inmediatamente inferior". Para ir entrando en calor, ¿por qué no prueba a calcular el valor de (sorpresa! 100)? (Si esta mañana ha tomado su ración cotidiana de fosfolípidos, hasta podría intentar hacerlo mentalmente.)

Sucede que a veces los Lisp-genios murmuran en voz alta mientras trabajan, y por casualidad yo pude entreoír este monólogo mientras el genio procedía a calcular "(sorpresa! 100)". Su soliloquio decía más o menos así:

"Humm... así que (sorpresa! 100), ¿eh? Bueno, sin duda 100 es distinto de 0, así que supongo que la solución habría de ser 100 veces la que hubiera sido si el problema consistiera en calcular (sorpresa! 99). Perfectamente, todo cuanto necesito ahora es conocer el valor de (sorpresa! 99). ¡Va a estar chupado! Veamos, ¿es 99 igual a 0? No, no parece serlo, así que la solución de *este* problema ha de ser 99 veces la correspondiente al caso (sorpresa! 98). Veamos ahora..."

En este momento, el autor, urgido por un negocio inaplazable, tuvo que dejar solo al genio feliz, y no pudo regresar hasta unas milésimas de segundo más tarde. Cuando llegó, el genio estaba a punto de terminar, diciendo:

"Y ahora me basta multiplicar *eso* por 100 y ya tendré la solución final. ¡Fácil como andar! Me parece que el resultado es 93326215443944152681699238856266700490715968264381621468592963895217599993229915608941463976156518286253697920827223758251185210916864000000000000000000000000, salvo que me haya equivocado".

¿Fue ese el resultado que usted obtuvo, querido lector? ¿No? ¡Ah, ya veo dónde se ha equivocado usted! Fue al multiplicar por 52. Repase desde allí, y ponga un poco más de atención al sumar esas largas columnas de números. Estoy seguro de que esta vez lo hará bien.

Esta función "sorpresa!" se llama, de ordinario, "factorial". La factorial de *n* suele definirse como el producto de todos los números naturales desde 1 hasta *n*. Empero, las definiciones recursivas miran el problema desde un ángulo ligeramente distinto: recursivamente hablando, la factorial de *n* no es más que el producto de *n* por la factorial precedente, lo que reduce así el problema a otro más sencillo del mismo tipo. Este, más sencillo, será a su vez reducido un paso más, y así una y otra vez, hasta llegar al problema más sencillo de la familia, que yo he venido llamando caso "embrionario" o "línea de base". En efecto, ya se utiliza la expresión "la recurrencia sale por la base" en la jerga del oficio.

Una viñeta del *New Yorker* de hace unos cuantos años ilustra la situación perfectamente. Se veía en ella un hombre de unos 50 años sosteniendo una fotografía de sí mismo tomada unos 10 años antes. En esa fotografía aparecía con otra fotografía suya de unos 10 años atrás, y así proseguía la escena hasta –literalmente– salirse por la base, en forma de fotografía de un nene retozón, vestido con el traje en que llegó al mundo (trasero al aire). La idea de una sucesión recurrente de fotografías que vayan captando como uno va creciendo es muy atractiva; ¡¿Cómo no se les ocurrió a mis padres!? Comparémosla con la regresión infinita de aquellos botes de leche condensada donde aparecía una granjera sosteniendo un bote de leche condensada con una granjera sosteniendo... Al no cambiar la edad de la moza, la regresión, en teoría, carecía de fin.

El enfoque recursivo va bien cuando se tiene una familia de problemas interrelacionados, de los cuales al menos uno es tan sencillo que podemos resolverlo inmediatamente: el que yo he llamado "caso embrionario". (En el ejemplo de la factorial, ese es el caso "(eq n 0)", cuya solución es 1.) Cada problema (por ejemplo, "¿cuánto vale factorial de 100?") puede considerarse como caso particular de un problema general ("¿cómo se calculan factoriales?"). La recurrencia saca partido del hecho de que las soluciones a los distintos casos están relacionadas unas con otras de cierta forma lógica. (Por ejemplo, yo podría decirle fácilmente el valor de factorial de 100 si usted me diera el de factorial de 99; todo cuanto he de hacer es multiplicar por 100 ese valor.) Podría decirse que el lema del buen "recursivador" es, "Bueno, yo sabría resolver *este* caso si alguien, por arte de magia, me facilitase la solución del caso

```
—> (montón-de-piedras 2)
    ENTERING montón-de-piedras (n = 2)
      ENTERING montón-de-piedras (n = 1)
        ENTERING montón-de-piedras (n = 0)
        EXITING montón-de-piedras    (valor: nil)
      EXITING montón-de-piedras    (valor: (piedra))
    EXITING montón-de-piedras    (valor: (piedra piedra))
(piedra piedra)
—>
```

2. El genio reúne un pequeño montón de piedras

situado un paso más cerca del embrión". Evidentemente, este lema presupone que ciertos casos están más cercanos al embrión que otros. De hecho, presupone que hay un sendero natural que va descendiendo desde un caso cualquiera hasta el caso embrionario a través de casos más sencillos; sendero, además, cuyos peldaños están marcados con claridad.

Resulta que tal presunción es de lo más razonable en todo tipo de circunstancias. Para captar claramente la naturaleza exacta de este sendero que guía la recurrencia, hay que responder a estas dos Grandes Preguntas:

1. ¿Cuál es el caso embrionario?
2. ¿Qué relación puede establecerse entre un caso típico y el caso inmediatamente más sencillo?

Ahora bien, en realidad estas dos Grandes Preguntas se escinden ambas en dos subcuestiones (como corresponde a toda cuestión recursiva que se respete a sí misma); a saber, una referente a cómo se reconoce dónde se está o cómo se ha de mover, y otra concerniente a cuál es la solución en cada etapa dada. Así pues, enunciadas más explícitamente, nuestras Grandes Preguntas son:

1*a*. ¿Cómo puede saberse que se ha alcanzado el caso embrionario?

1*b*. ¿Cuál es la solución embrionaria?

2*a*. Partiendo de un caso típico, ¿cómo hacer, exactamente, para dar un paso en dirección a la solución embrionaria?

2*b*. ¿Cómo se construiría la solución del caso típico si "por arte de magia" se conociera la solución del caso inmediatamente más sencillo?

La cuestión 2*a* concierne a la naturaleza del *descenso* hacia el caso embrionario, hacia la línea de base. La cuestión 2*b* trata del aspecto inverso, a saber, el *ascenso* que nos lleva de regreso desde la línea de fondo hasta el nivel de superficie.

En el caso de la factorial, las soluciones a las Grandes Preguntas son:

1*a*. Se da el caso embrionario cuando el argumento es 0.

1*b*. La solución al caso embrionario es 1.

2*a*. Réstese 1 del valor que actualmente tenga el argumento.

```
—> (ruta-hacia-1 3)
ENTERING ruta-hacia-1 (aysa = 3)
  ENTERING ruta-hacia-1 (aysa = 10)
    ENTERING ruta-hacia-1 (aysa = 5)
      ENTERING ruta-hacia-1 (aysa = 16)
        ENTERING ruta-hacia-1 (aysa = 8)
          ENTERING ruta-hacia-1 (aysa = 4)
            ENTERING ruta-hacia-1 (aysa = 2)
              ENTERING ruta-hacia-1 (aysa = 1)
              EXITING ruta-hacia-1 (valor: (1))
            EXITING ruta-hacia-1 (valor: (2 1))
          EXITING ruta-hacia-1 (valor: (4 2 1))
        EXITING ruta-hacia-1 (valor: (8 4 2 1))
      EXITING ruta-hacia-1 (valor: (16 8 4 2 1))
    EXITING ruta-hacia-1 (valor: (5 16 8 4 2 1))
  EXITING ruta-hacia-1 (valor: (10 5 16 8 4 2 1))
EXITING ruta-hacia-1 (valor: (3 10 5 16 8 4 2 1))
(3 10 5 16 8 4 2 1)
—>
```

3. Así "razona" el genio cuando está activado el "trazador"

2*b*. Multiplíquese la solución "mágica" por el valor actual del argumento.

Observemos que las respuestas a estas cuatro preguntas están todas incorporadas a la definición recursiva de "¡sorpresa!".

La recurrencia descansa en la hipótesis de que antes o después "saldremos por la base". Una forma de estar *seguro* de salir por la base es que los pasos "descendentes", o de "simplificación", procedan todos en la misma dirección y al mismo ritmo, con lo que nuestra trayectoria será, evidentemente, rectilínea. Por ejemplo, es obvio que restando 1 reiteradamente acabaremos llegando a 0 si partimos de un entero positivo. Es igualmente obvio que efectuando la operación de abreviación de listas llamada "cdr" se acabará alcanzando nil, suponiendo, claro está, que se haya comenzado con una lista finita. Por esta razón, las recurrencias que utilizan "cdr" o "sub1" para definir un sendero descendente son corrientísimas. Más abajo presentaré una recurrencia basada en cdr, pero antes deseo proponerles una entretenida recurrencia numérica, donde la senda hacia el caso embrionario tiene tan poco de regular como de rectilínea.

Fijémonos en el famoso problema "$3n + 1$", donde se comienza con un entero positivo cualquiera y, si éste es par, se le divide entre dos; en caso de ser impar, se le multiplica por tres y se le suma 1 al producto. El resultado de someter n a esta operación será llamado "(hotpo n)" ("half, or triple plus one", "la mitad, o el triple más uno"). He aquí una definición de hotpo:

```
(def hotpo
  (lambda (n)
    (cond ((par n) (mitad n))
          (t (suma 1 (producto 3 n))))))
```

Esta definición presupone que otras dos funciones han sido ya –o van a ser– definidas en algún otro lugar para información del Lisp-genio, a saber, las funciones "par" y "mitad". ("Suma 1" y "producto" son, como ya se ha explicado, partes intrínsecas del Lisp.) He aquí las definiciones que faltan:

```
(def par
 (lambda (n) (eq (resto-división n 2) 0)))

(def mitad (lambda (n) (cociente n 2)))
```

¿Qué piensa usted que ocurrirá si se toma un entero cualquiera y se efectúa sobre él hotpo una y otra vez? Pongamos, por ejemplo, que el valor de partida sea 7. Antes de hacer los cálculos aritméticos, trate de conjeturar qué fenómeno podremos observar.

Resulta que la trayectoria es con frecuencia tan caótica como accidentada. Por ejemplo, si comenzamos con 7, el proceso nos conducirá a 22, después a 11, seguidamente a 34, a 17, a 52, 26, 13, 40, 20, 10, 5, 16, 8, 4, 2, 1, 4, 2, 1, 4, 2, 1,... Observemos que al final caemos en un bucle de período corto. Supongamos, por otra parte, que convenimos en que al llegar a 1 tocamos fondo y po-

demos parar. Podríamos muy bien preguntar: "¿Y quién nos garantiza que llegaremos alguna vez a 1? ¿Hay seguridad en ello?" Y en efecto, si no se ha ensayado un poco, no salta a la vista que vayamos a alcanzar el 1. (¿Sospechó usted que así llegaría a ocurrir en el caso del 7?) No obstante, al experimentar con números se aprecia una notable regularidad en el proceso; parece que, con independencia de dónde se comience, siempre se acaba dando con el 1 al cabo de algún tiempo. (Si quiere darse una vuelta en las montañas rusas, pruebe a empezar con 27.)

¿Podrá el lector construir una función recursiva que vaya mostrando la senda "descendente" seguida desde un punto de partida cualquiera hasta 1? Adviértase que digo "descendente" con premeditación, pues en realidad muchos de los pasos son *ascendentes*. Así, la senda que parte de 3 será la lista "(3 10 5 16 8 4 2 1)". Para poder resolver este problema es necesario retroceder, y contestarnos a nosotros mismos las dos grandes cuestiones de la recurrencia, en la forma aquí pertinente. Observe:

```
(cond ((no (necesito ayuda))
        (no (prosiga leyendo)))
      (t (prosiga leyendo)))
```

Empecemos examinando el caso embrionario. Es cosa fácil. Ha sido definido ya como llegada al 1, y la solución embrionaria, la de máxima sencillez, es la lista "(1)", que es una breve, aunque legítima, senda, que empieza y acaba en 1.

Seguidamente, veamos casos más típicos. ¿Qué operación nos acercará un paso al nivel básico si partimos de un valor típico, como 7? No va a ser, desde luego, la operación "sub1". No; por definición es la propia función hotpo la que cada vez nos lleva "más cerca" de 1, incluso cuando nos conduce hacia valores *altos*. Esta molesta cualidad es la que da interés al ejemplo. ¿Qué diremos de 2*b*, que pide cómo construir recursivamente una lista que documente nuestra abrupta y serpeante senda? Bueno, la senda correspondiente al 7 se obtiene adosando (es decir, consando) un 7 a la senda inmediatamente más breve, asociada ésta con (hotpo 7), es decir, 22. Después de todo, 22 está un paso más cerca de ser embrión que lo está el 7.

Estas consideraciones nos permiten escribir la definición de la función requerida, valiéndonos de "aysa" como variable ficticia. (Como se sabe, "aysa" es acrónimo de "aysa (y solamente aysa)", que por recurrencia se expande a "aysa (y solamente aysa) (y solamente (y solamente aysa) aysa)", y así progresivamente.

```
(def ruta-hacia-1
 (lambda (aysa)
  (cond ((eq aysa 1) '(1))
        (t
         (cons aysa
           (ruta-hacia-1
             (hotpo aysa)))))))
```

Eche un vistazo a la forma de "razonar" del Lisp-genio (tal como nos lo muestra el dispositivo "trazador" al hacerlo funcionar), que podemos ver en la figura 3.

Notemos la perfecta regularidad (en forma de *V*) del margen izquierdo del diagrama de la traza, a pesar de lo caótico de los números implicados. No todas las recurrencias presentan tan gran regularidad geométrica. La razón está en que ciertos problemas exigen resolver *más de un subproblema*. Un ejemplo de este tipo, práctico y tomado de la vida real, nos lo proporciona el problema de los unicornios aún supervivientes en Europa. No es desde luego una empresa trivial, pero hay, no obstante, una hermosa y elegante solución por recurrencia: cuéntese todos los unicornios residentes en Portugal (que es, metafóricamente hablando, la "car" de Europa) y céncese después todos los unicornios de los otros 30 países europeos (la "cdr" de Europa, por machacar en la metáfora) y, finalmente, súmese ambos resultados parciales.

Fijémonos en que se generan dos subproblemas de recuento de unicornios, los cuales darán origen a dos subproblemas cada uno, siempre de menor tamaño; y así sucesivamente. Por ejemplo, ¿cómo podemos contar todos los unicornios de Portugal? Es cosa fácil: se suman los que haya en la Extremadura portuguesa (la "car" de Portugal) con el número de unicornios del resto del país ("cdr" de Portugal). ¿Y cómo se cuentan los unicornios de Extremadura, por no mentar los de otras regiones? Pues por descomposición adicional, desde luego. ¿Y cuándo "tocaremos fondo", así pues? Bueno, las regiones se descomponen en distritos, los distritos en kilómetros cuadrados; éstos, en hectáreas, las hectáreas en metros cuadrados –y para buscar unicornios, ya no es preciso dividir más finamente.

Aunque todo esto puede parecer auténticamente arduo, la verdad es que no hay más forma de elaborar un censo concienzudo que ir explorando unidad por unidad cada uno de los niveles de la estructura que se tenga, por gigantesca que pueda ser. Tropezamos aquí con un problema que es perfectamente equivalente en Lisp al de censar los unicornios, a saber, el problema de determinar cuántos átomos hay en una lista arbitraria. ¿Cómo podemos escribir una función en Lisp –a la que llamaremos "recuento-atómico"– que nos proporcione la respuesta 15 cuando le sea presentada la siguiente lista, de exótica apariencia, a la que llamaremos "brahma"?

```
(((ac ab cb) ac (ba bc ac))
               ab
((cb ca ba) cb (ac ab cb)))
```

Un posible método es, recursivamente expresado, homólogo exacto del explicado para conocer la población de unicornios de Europa. A ver si puede usted redactarla.

He aquí la idea. Deseamos construir la respuesta –a saber, 15– a partir de las soluciones a problemas más sencillos de recuento atómico. Es evidente que un problema más sencillo, de igual tipo que (recuento-atómico brahma) es (recuento-atómico (car brahma)). Otro es (recuento (cdr brahma)). Las respuestas de estos dos subproblemas son, respectivamente, 7 y 8. Ahora, como es obvio, 15 resulta de sumar 7 más 8 –lo cual bien era de esperar, pues el número total de átomos ha de ser el que contenga la sección car más el que corresponda a la sección cdr. No hay ningún otro sitio donde puedan los átomos esconderse. Este análisis nos da la siguiente definición recursiva, donde "s" es la variable muda:

```
(def recatom
 (lambda (s)
  (suma (recatom (car s))
        (recatom (cdr s)))))
```

Aunque la definición parezca sencilla, tiene un par de fallos. Primero, hemos escrito la parte *recursiva* de la definición, pero hemos olvidado otra mitad no menos importante: la línea de fondo. Lo cual me trae a las mientes el caso de un juez de Maryland quien, según leí en un periódico, dictaminó que "un caballo es un animal cuadrúpedo, que es engendrado por dos caballos". Como definición es encantadora, pero ¿cómo se sale por la línea de base? Otro tanto le sucede a nuestro

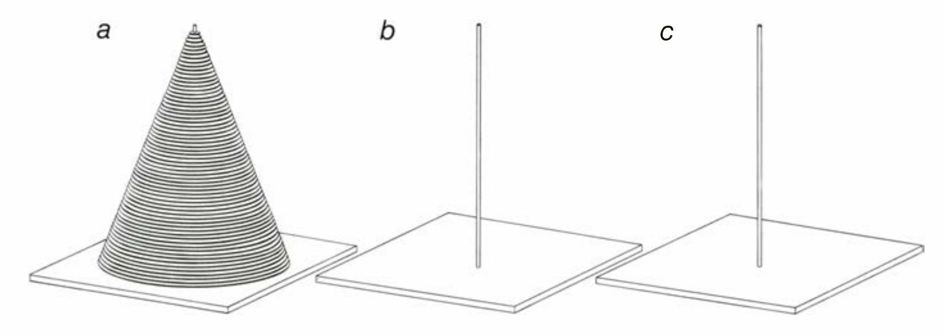
"recuento-atómico". ¿Cuál es, en él, el caso más sencillo, o caso embrionario? Se produce cuando se nos pide contar los átomos de que consta un átomo individual. La respuesta es, en tal caso, 1, evidentemente. ¿Pero cómo podemos saber que estamos habiéndonoslas con un átomo individual? Afortunadamente hay en Lisp una función intrínseca, llamada "átomo", que retorna t (es decir, "true", o verdadero) cada vez que damos con un átomo y nil en los restantes casos. Así pues, "(átomo 'plop)" genera el retorno t; "(átomo '(a b c))" retorna nil. Gracias a esto podemos remendar nuestra definición así:

```
(def recatom
 (lambda (s)
  (cond ((átomo s) 1)
      (t
       (suma (recatom (car s))
             (recatom (cdr s)))))))
```

Empero, la definición no es correcta del todo. Si pedimos al genio el recuento de "(a b c)", en lugar de darnos 3, nos da como respuesta 4. ¡Escandaloso! ¿Cómo habrá podido ocurrir? Para mejor detectar la raíz del problema, examinemos un caso aún más sencillo: si pedimos (recatom '(a)), en lugar de 1 recibimos un 2. Ahora debería estar más claro cuál es el error: 2 = 1 + 1, donde un 1 procede de la car y otro de la cdr de "(a)". La car es el átomo "(a)", que debe ser contado como 1,

```
—> (recatom brahma)
    ENTERING recatom
       (s = (((ac ab cb) ac (ba bc ac)) ab ((cb ca ba) cb (ac ab cb))))
      ENTERING recatom (s = ((ac ab cb) ac (ba bc ac)))
        ENTERING recatom (s = (ac ab cb))
          ENTERING recatom (s = ac)
          EXITING    recatom (valor: 1)
          ENTERING recatom (s = (ab cb))
            ENTERING recatom (s = ab)
            EXITING    recatom (valor: 1)
            ENTERING recatom (s = (cb))
              ENTERING recatom (s = cb)
              EXITING    recatom (valor: 1)
              ENTERING recatom (s = nil)
              EXITING    recatom (valor: 0)
            EXITING recatom (valor: 1)
          EXITING recatom (valor: 2)
        EXITING    recatom (valor: 3)
        ENTERING recatom (s = (ac (ba bc ac)))
          ENTERING recatom (s = ac)
          EXITING    recatom (valor: 1)
          ENTERING recatom (s = ((ba bc ac)))
            ENTERING recatom (s = (ba bc ac))
              ENTERING recatom (s = ba)
              EXITING    recatom (valor: 1)
              ENTERING recatom (s = (bc ac))
                ENTERING recatom (s = bc)
                EXITING    recatom (valor: 1)
                ENTERING recatom (s = (ac))
                  ENTERING recatom (s = ac)
                  EXITING    recatom (valor: 1)
                  ENTERING recatom (s = nil)
                  EXITING    recatom (valor: 0)
                EXITING recatom (valor: 1)
              EXITING recatom (valor: 2)
            EXITING    recatom (valor: 3)
            ENTERING recatom (s = nil)
            EXITING    recatom (valor: 0)
          EXITING recatom (valor: 3)
        EXITING recatom (valor: 4)
      EXITING    recatom (valor: 7)
      ENTERING recatom
         (s = (ab ((cb ca ba) cb (ac ab cb))))
        ENTERING recatom (s = ab)
        EXITING    recatom (valor: 1)
        ENTERING recatom
           (s = (((cb ca ba) cb (ac ab cb))))
          ENTERING recatom (s = ((cb ca ba) cb (ac ab cb)))
            ENTERING recatom (s = (cb ca ba))
              ENTERING recatom (s = cb)
              EXITING    recatom (valor: 1)
              ENTERING recatom (s = (ca ba))
                ENTERING recatom (s = ca)
                EXITING    recatom (valor: 1)
                ENTERING recatom (s = (ba))
                  ENTERING recatom (s = ba)
                  EXITING    recatom (valor: 1)
                  ENTERING recatom (s = nil)
                  EXITING    recatom (valor: 0)
                EXITING recatom (valor: 1)
              EXITING recatom (valor: 2)
            EXITING    recatom (valor: 3)
            ENTERING recatom (s = (cb (ac ab cb)))
              ENTERING recatom (s = cb)
              EXITING    recatom (valor: 1)
              ENTERING recatom (s = ((ac ab cb)))
                ENTERING recatom (s = (ac ab cb))
                  ENTERING recatom (s = ac)
                  EXITING    recatom (valor: 1)
                  ENTERING recatom (s = (ab cb))
                    ENTERING recatom (s = ab)
                    EXITING    recatom (valor: 1)
                    ENTERING recatom (s = (cb))
                      ENTERING recatom (s = cb)
                      EXITING    recatom (valor: 1)
                      ENTERING recatom (s = nil)
                      EXITING    recatom (valor: 0)
                    EXITING recatom (valor: 1)
                  EXITING recatom (valor: 2)
                EXITING    recatom (valor: 3)
                ENTERING recatom (s = nil)
                EXITING    recatom (valor: 0)
              EXITING recatom (valor: 3)
            EXITING recatom (valor: 4)
          EXITING    recatom (valor: 7)
          ENTERING recatom (s = nil)
          EXITING recatom (valor: 0)
        EXITING recatom (valor: 7)
      EXITING recatom (valor: 8)
    EXITING recatom (valor: 15)
15
—>
```

4. Trazado de la función Lisp "recatom" al aplicarla a la lista "brahma"



5. El rompecabezas de la Torre de Brahma.
Los 64 discos enclavados en "a" deben quedar en "b", en idéntico orden

mientras que la cdr es nil, que no debería serlo. ¿Por qué, pues, da nil un recuento de 1? ¡Porque nil no sólo es una lista, sino también un átomo! Para eliminar este nocivo efecto secundario insertaremos otra cláusula condicional al principio de nuestra definición.

```
(def recatom
 (lambda (s)
  (cond ((nulidad s) 0)
        ((átomo s) 1)
        (t
         (suma (recatom (car s))
               (recatom (cdr s)))))))
```

He escrito "(nulidad s)", que no es sino otra forma de decir "(eq s nil)". En general, cuando se desea determinar si el valor de cierta expresión es, o no, igual a nil, puede utilizarse la función intrínseca "nulidad", que retorna t en caso afirmativo, y nil en el negativo. Así, por ejemplo, "(nulidad (nulidad nil))" produce nil al ser evaluada, pues la función del paréntesis interno produce t, y t no es nil! ¿Qué pasará con "(nulidad '(nulidad nil))"?

Veamos ahora qué sucederá cuando hagamos actuar nuestra función sobre la lista brahma, que era su objetivo inicial. El resultado, con la traza en funcionamiento, aparece en la figura 4.

Fijémonos en la mucho más complicada topografía de esta recurrencia, en los entrantes y salientes que muestra a lo largo del margen izquierdo. La recurrencia anterior, con su *V* perfecta, parecía un sencillo descenso a un valle, seguido de una ascensión igualmente sencilla por la otra ladera; la ruta que ahora nos ocupa es mucho más accidentada, como si en el ascenso y en el descenso por cada pared encontrásemos distintos valles y quebradas, que debemos tratar de igual manera –y quién sabe con cuántos niveles de semejante estructura habremos de tropezar en nuestra exploración. Las formas descritas por estructuras que prosiguen indefinidamente a la manera de ésta han sido bautizadas con el nombre de "fractales" por Benoit Mandelbrot.

Observemos también que en esta recurrencia encontramos más de un tipo de caso embrionario (el caso de "nulidad", y el caso "atómico") y también, con más de un procedimiento para descender hacia el caso embrionario (mediante car y cdr). Así pues, nuestras Grandes Preguntas pueden revisarse un poco más profundamente:

1*a*. ¿Hay únicamente un caso embrionario, hay varios o hay toda una clase infinita de posibles casos "elementales"?

1*b*. ¿De qué forma podemos saber que hemos alcanzado un caso embrionario?

1*c*. ¿Qué soluciones tienen los distintos casos embrionarios?

2*a*. A partir de un caso típico, ¿hay exactamente una forma de avanzar hacia el caso embrionario, o hay varias?

2*b*. Dado un caso típico, ¿cómo se determina la vía a tomar hacia un caso embrionario?

2*c*. ¿Cómo se construye la solución de este caso típico a partir de soluciones "mágicamente dadas" a uno o varios casos más sencillos?

Una de las recurrencias más elegantes que conozco se origina en el famoso rompecabezas conocido con los nombres de "Torre de Lucas", "Torre de Hanoi" y "Torre de Brahma". Según parece, este rompecabezas surgió de la imaginación de Édouard Lucas, un matemático francés del siglo pasado. He aquí la leyenda popular atribuida al rompecabezas:

En el gran templo de Brahma, en Benarés, sobre una placa de bronce situada bajo el domo que define el centro del mundo, hay 64 discos perforados, de oro puro, que los sacerdotes del templo van trasladando a razón de uno por vez entre tres finas agujas de diamante, obedeciendo a esta inmutable ley de Brahma: Nunca puede colocarse un disco sobre otro de diámetro menor. Al principio del mundo los 64 discos formaban en torno a la aguja la Torre de Brahma. En la actualidad, el proceso de traslado se halla por su mitad. Cuando finalmente el último disco se sitúe en lugar, formando de nuevo la Torre de Brahma en torno a otra aguja diferente, llegará el fin del mundo y todo se reducirá a polvo.

Vemos en la figura 5 un dibujo del rompecabezas; he rotulado en él "a", "b" y "c" para distinguir las tres agujas.

Si el lector dedica un momento al problema descubrirá el procedimiento sistemático que han de seguir los brahmanes para ir transfiriendo los discos desde la varilla "a" hasta la varilla "b". Por ejemplo, con sólo tres discos no es difícil expresar el orden que han de seguir los movimientos:

ab ac bc ab ca cb ab

Representamos aquí el paso de un disco desde la varilla "a" a la "b" por un átomo, "ab" en Lisp. Hay, empero, una estructura en lo que va sucediendo que no queda manifiesta por el mero listado de tales átomos. Más clara quedará agrupándolos como sigue:

ab ac bc   ab   ca cb ac

El primer grupo de tres efectúa la transferencia de una torre de altura 2 desde la varilla "a" hasta la "c", dejando así accesible el disco mayor. Seguidamente, en el movimiento central, "ab", se toma este disco grande y pesado, y se lo traslada desde la aguja "a" hasta la "b". El grupo final, compuesto también por tres átomos, es muy parecido al trío inicial, pues transfiere la torre de altura 2 desde la aguja "c" hasta la aguja "b". Por tanto, resolver el problema de trasladar tres discos depende de que pueda resolverse el de trasladar dos. Análogamente, la solución del problema de los 64 discos pasa por la del problema para 63. ¿Será preciso decir más? Intente ahora el lector redactar en Lisp la definición de una función que resuelva el problema de la Torre de Brahma para *n* discos. (Tal vez le convenga marcar las varillas con números en lugar de letras, con lo que los movimientos quedarán representados por números de dos cifras, como 12.) Presentaré la solución el mes que viene, siempre, claro está, que los entregados sacerdotes, que laboran noche y día para hacernos llegar el fin del mundo, no tengan la suerte de alcanzar tan ambicionada meta antes de entonces.

Taller y laboratorio

Cómo examinar, desde fuera, un sistema de semáforos de su ciudad

Jearl Walker

El automovilista que conduce por la ciudad sabe que, tarde o temprano, llegará un momento en que los semáforos le obligarán a parar en casi todos los cruces. A partir de ese instante, su marcha se hará morosa y le invadirá una sensación de pérdida de tiempo inmerso en un embotellamiento de proporciones más o menos imponentes. Entonces, pensará que las señales que encuentra a su paso deberían estar sincronizadas y así dar fluidez al tráfico.

Lo cierto es que, en la mayoría de las ciudades, las rutas de tránsito importantes disponen ya de sistemas de semáforos sincronizados, especialmente durante las horas punta. Estos sistemas se basan en la hipótesis de que todo conductor que viaje a una velocidad determinada ha de encontrar una luz verde en cada cruce. Cuando el tránsito es denso, tal temporización obliga a los automóviles a agruparse en "pelotones", por emplear un término acuñado por los analistas de circulación. En los intervalos entre pelotones, el tránsito de las calles perpendiculares a la principal se cruza con el de ésta o se mezcla con él. En un trayecto concebido correctamente, la secuencia de las señales debe hacer que la circulación se desarrolle sin retrasar excesivamente el tránsito perpendicular.

Con el ánimo de examinar un sistema de señales sincronizadas fijé mi atención en los semáforos de una calle céntrica de mi ciudad, Cleveland. Seguramente me hubiera sido posible averiguar todo lo que deseaba preguntando en el ayuntamiento, pero no me habría divertido tanto. Además, hay siempre la posibilidad de que un analista independiente sugiera alguna idea aprovechable. Un estudio de este tipo puede hacerlo cualquiera en su propia ciudad.

Los semáforos de Carnegie, que así se llama la calle escogida, están gobernados por un sistema electrónico cerrado y no por un sistema programado por la densidad de la circulación. Esta calle es una arteria principal que une el centro de Cleveland con los extrarradios del este y soporta un tránsito densísimo durante las horas punta de las mañanas y del atardecer. Durante la hora punta de la mañana el tránsito que se dirige hacia el oeste dispone de cuatro carriles, y de dos el que se dirige hacia el este; a la hora punta del atardecer todos los carriles se dirigen hacia el este. El resto del día hay tres carriles para cada sentido.

La calle Carnegie atraviesa una zona donde se concentran el Teatro de Cleveland y múltiples tiendas y almacenes, amén de viviendas. Las calles laterales son todas de tres carriles y enlazan Carnegie con otras vías paralelas a ella, por las cuales se canaliza también el tránsito de las horas punta. Entre las calles principales discurre un tráfico considerable, por lo que las señales rojas de las calles laterales no pueden demorarse mucho, ya que se provocarían colas que terminarían taponando las avenidas principales.

Me ocupé del sistema de señales que regulaba la calle Carnegie durante los períodos iniciales de la hora punta del atardecer. Partiendo del cruce de las calles Carnegie y 71-Este, medí la distancia desde cada cruce hasta el siguiente caminando por Carnegie en el mismo sentido que la circulación. En todos los cruces cronometré la duración de las luces roja, ámbar y verde. Además, a la vez que realizaba el recorrido, tomé nota del intervalo entre el inicio de la señal verde en cada cruce y el inicio de la misma señal en el siguiente.

Como útil de trabajo me serví de un reloj digital dotado de cronómetro y de lectura de tiempo desdoblada, especialmente práctico. Cuando en un cruce la luz se ponía ámbar activaba el cronómetro y, cuando se ponía roja o verde, oprimía el botón de lectura desdoblada. Este dispositivo retenía unos instantes la indicación de la esfera, lo que me permitía anotar el tiempo en mi cuaderno.

El cronómetro, empero, proseguía su marcha y, al cabo de cinco segundos, la esfera registraba, como antes, el paso de los segundos.

Tras numerosas observaciones saqué la media de las lecturas correspondientes a cada color para reducir el error debido a las variaciones del tiempo que tardaba en reaccionar ante los cambios de luz y pulsar los botones del reloj. También promedié mis medidas del tiempo que transcurría entre la aparición de la luz verde en un cruce y la aparición en el siguiente.

Las distancias entre los cruces las medí contando pasos entre las líneas de parada de cada dos cruces sucesivos. De regreso a casa repetí aproximadamente los mismos pasos en el patio. Midiendo esas distancias con una vara graduada convertí a metros las longitudes que había medido a lo largo de Carnegie. Para reducir el error que pudieran producir las variaciones en la longitud de mi zancada, en mi casa caminé siempre todo el ancho del patio. Aún así, imagino que el error que cometí al medir la distancia entre cruces se encuentra entre el 10 y el 20 por ciento. Desde luego, una medición directa hubiera sido mucho más exacta, pero la ganancia en precisión no hubiera compensado el esfuerzo adicional.

En la figura 1 se resumen estas medidas. Abajo, a la izquierda, se representa el cruce de Carnegie y 71-Este. En el dibujo de la calle Carnegie la circulación tiene sentido ascendente; a lo largo de la misma aparecen los demás cruces.

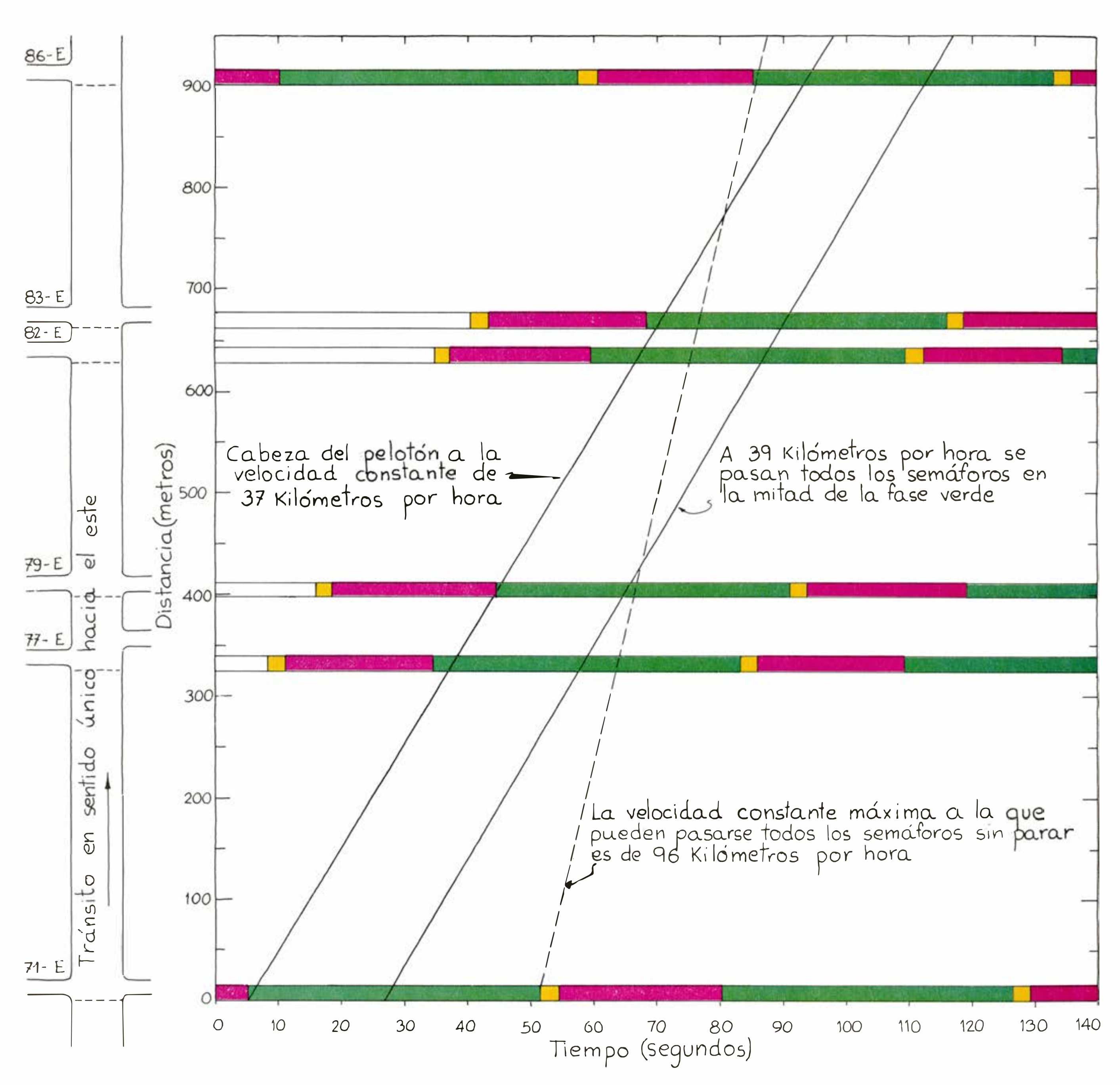
En la escala horizontal de la gráfica indico los tiempos de las luces. Por ejemplo, la luz verde en Carnegie y 71-Este se representa iniciándose en el instante cinco segundos. Este fue el punto de referencia que me sirvió para medir esa luz y la de los cruces sucesivos. Como la duración media de la luz verde en 71-Este era de 46,4 segundos, en la gráfica se señala el comienzo de la luz ámbar a los 51,4 segundos, suma de 46,4

segundos más los cinco segundos de referencia. Análogamente, se recogen el principio de la luz roja y los instantes correspondientes al ciclo siguiente de verde, ámbar y rojo.

En el cruce siguiente, que es el de Carnegie con 77-Este, la luz se pone verde 29,5 segundos después de que lo haga en 71-Este. Por tanto, en la gráfica el inicio de la fase verde en 77-Este se señala con 34,5 segundos, que es la suma de 29,5 segundos y los cinco segundos de referencia. Al durar la luz verde en ese punto 48,6 segundos, la aparición de la luz ámbar la indica la gráfica con 83,1 segundos. Los ciclos de colores en los demás cruces se representan análogamente respecto a la referencia de cinco segundos añadidos a la luz verde de la calle 71-Este.

Durante la hora punta del atardecer se forman pelotones de vehículos en la zona de Carnegie contigua a 71-Este, o más cercana al centro de la ciudad. Los pelotones avanzan hacia el este a velocidad casi constante. Un cabeza de pelotón podría apresurarse para ir de cruce a cruce, pero de poco le serviría: durante un tiempo y a lo largo de una distancia extensos no es posible ir más deprisa de lo que permite la secuencia de las luces. Si un cabeza parte de un cruce a la máxima aceleración y a velocidad excesiva, el automóvil llegará al cruce siguiente con demasiada antelación sobre la luz verde.

En la gráfica se representa el movimiento de un cabeza de pelotón que haga el recorrido de los seis cruces a velocidad constante. Este automóvil parte cuando se pone verde el semáforo de la calle 71-Este, realizando el trayecto a velocidad constante para no toparse con rojos. Según mis mediciones, las señales no están perfectamente sincronizadas en lo que respecta al cabeza. Si el conductor ha de pasar por el cruce con la calle 79-Este precisamente cuando se encienda el verde, llegará a la 82-Este unos cinco segundos después de que en ésta aparezca ese color. Quizás



1. Sistema de señales de circulación en la calle Carnegie durante la hora punta del atardecer

77-E
Cabezas
Pelotón delantero
Ultimo vehículo
Velocidad del cabeza, 10 metros por segundo
71-E
Onda de puesta en marcha, 5 metros por segundo
Pelotón de cola
El semáforo de la calle 71 se pone verde
Velocidad del cabeza, 10 metros por segundo
Onda de puesta en marcha, 5 metros por segundo
Posición de la onda de puesta en marcha
15 segundos después
Detención de los cabezas de pelotón
Posición de la onda de puesta en marcha
El semáforo de la calle 77 se pone verde

2. Situaciones que pueden desembocar en un embotellamiento

este desajuste forme parte de la programación, pues ello permitiría que el tránsito que afluye a Carnegie por 79-Este despeje los cruces con 82-Este y 83-Este antes de la arribada del pelotón siguiente.

En la gráfica se representan asimismo otros dos conductores hipotéticos que recorran los seis cruces a velocidad constante. Uno de ellos los atraviesa aproximadamente a la mitad de la fase verde. El otro conductor pasa por 71-Este justo al final de la fase verde y luego acelera para cruzar por 86-Este en el momento en que la luz del semáforo se pone verde. En la gráfica pueden medirse las velocidades del cabeza de pelotón y de los otros dos conductores: la velocidad, en metros por segundo, es la pendiente de la recta que representa a cada uno de ellos.

El cabeza de pelotón viaja a 10,3 metros por segundo, unos 37 kilómetros por hora. El conductor que pasa por los semáforos hacia la mitad de la fase verde viaja a 10,8 metros por segundo, o sea, a 39 kph aproximadamente. El conductor que atraviesa el sistema a toda prisa en el tiempo mínimo se desplaza a 26,6 metros por segundo, unos 96 kph. (La velocidad límite en la calle Carnegie es de 40 kph.)

Considerando lo que sucede en el tránsito real, nuestros conductores imaginarios, tan respetuosos de la ley, avanzarían algo lentos. Cuando, a diario, conduzco camino de casa por el sistema de la hora punta, mi pelotón suele moverse a unos 45 kph. En la discrepancia entre esta velocidad y la calculada con la gráfica intervienen tres factores. Si la circulación ha de ser uniforme, un cabeza de pelotón debe disponer de luz verde uno o dos segundos antes de llegar a cada cruce; si no, por precaución aminorará la marcha. La gráfica, además, pasa por alto el tiempo que tarda el cabeza de pelotón en reaccionar tras señalarse la luz verde en la calle 71-Este. Ese tiempo puede ser del orden del segundo.

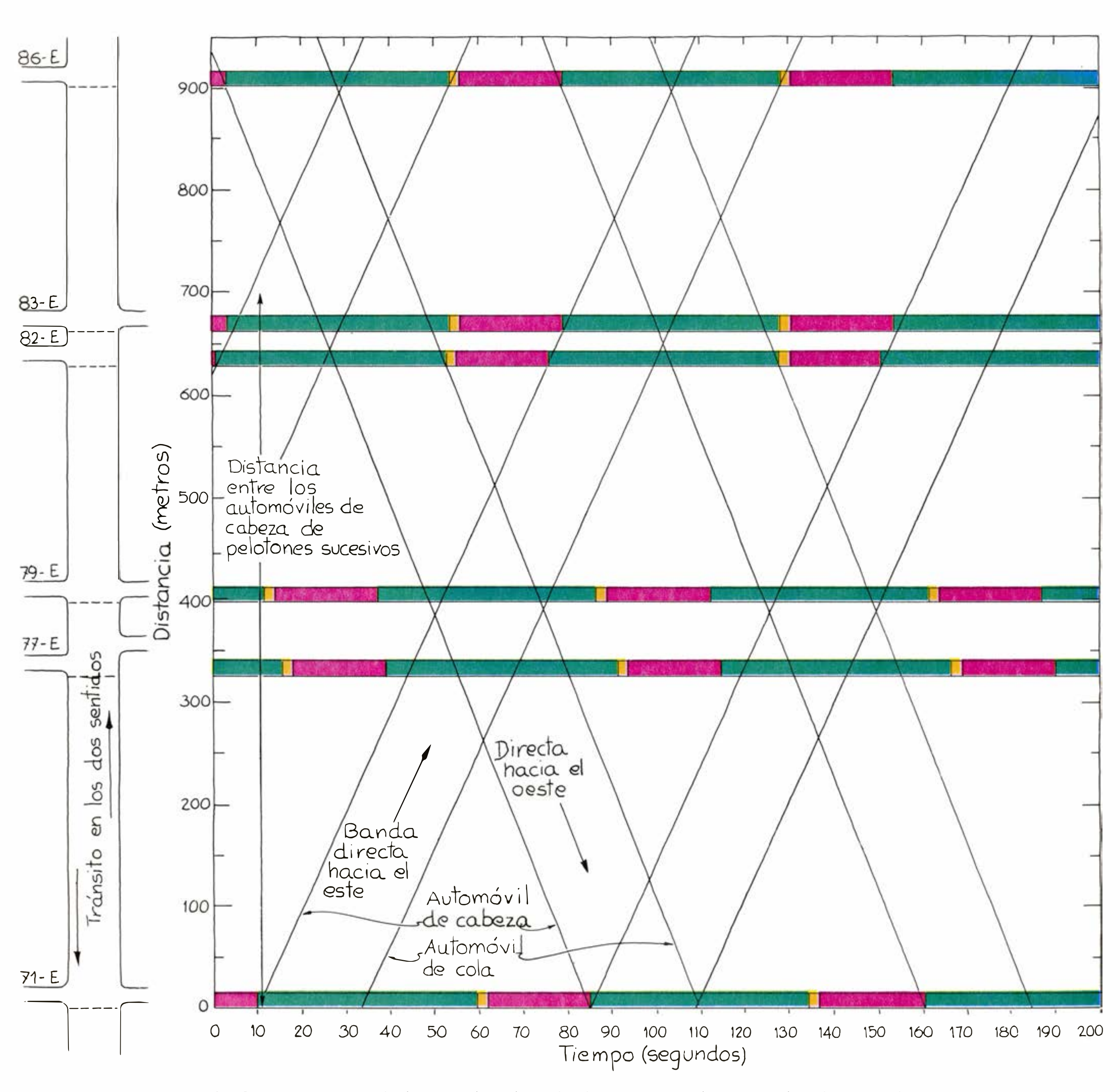
El tercer factor es que la gráfica no tiene en cuenta el tiempo necesario para acelerar un automóvil desde el reposo hasta la velocidad constante que adquiere el pelotón dentro del sistema. Cuando yo encabezo la fila, necesito unos tres segundos para alcanzar la velocidad de marcha. Estos tres factores reducen el tiempo de que dispone un cabeza de pelotón para ir de un cruce al siguiente. Por tanto, puede conducir a una velocidad constante algo superior a los 37 kph sin tener que detenerse en ningún cruce.

A veces he ajustado mi automóvil a ese sistema después de que Carnegie haya comenzado a ser calle de sentido único, pero antes de que haya crecido el volumen de circulación. Entonces, si paso por 71-Este al final de la fase verde, consigo recorrer algunos de los cruces a más de 80 kph, pero me detiene la luz roja en la calle 86-Este. Casi siempre que sobrepasé el límite de velocidad me adelantó otro conductor, por lo que deduzco que será posible viajar por ese tramo a una velocidad constante de más de 90 kph.

Durante las horas punta, circulan por Carnegie pelotones de vehículos cada 75 segundos, más o menos, que es el intervalo entre las sucesivas luces verdes en la calle 71-Este. Aunque las luces no están perfectamente sincronizadas a lo largo del recorrido, el sistema funciona suficientemente bien para que los pelotones no resulten excesivamente largos para las distancias que separan los cruces. De hecho, la fase verde podría rebajarse en 10 o 15 segundos en la mayoría de ellos sin retardar la marcha promedia. Yo diría que la duración de las fases verdes se mantiene así para el caso de que surja algún imprevisto. Ese tiempo adicional se necesitará si un accidente atora alguno de los carriles o si el tránsito de las horas punta adquiere cierta densidad.

La circulación por la calle Carnegie al comienzo de la hora punta del atardecer tiende a congregarse hacia la derecha, aunque muchos conductores evitan el primer carril de esta mano, dada la posibilidad de encontrar un automóvil averiado o estacionado ilegalmente. Pero, conforme aumenta la circulación, los carriles empiezan a llenarse más uniformemente.

Hasta ese momento, los pelotones pueden maniobrar para salvar cualquier obstáculo que haya en un carril. Cuando la circulación por un carril se hace más lenta, los automovilistas situados al final de la cola se pasan a los otros carriles. Las dificultades se pro-



3. Sincronización de las señales al mediodía para circulación en los dos sentidos

ducen cuando éstos ya están llenos. Entonces la cola que se forme tras un obstáculo puede desbaratar el programa ideado, a través de la red de señales, para aligerar la circulación.

El plan del sistema de señales en la calle Carnegie parece basarse en la hipótesis de que, cuando un cabeza de pelotón se acerque a un cruce, el pelotón precedente ya habrá salido. Pero si parte de este último sigue obstruyendo el paso, el cabeza del nuevo tendrá que reducir la velocidad o detenerse, variando de sentido la temporización de las luces. Ese retraso podría suponer que parte del nuevo pelotón no atraviese el cruce antes de la llegada del siguiente. Si así es, la situación empeorará cada vez más.

La figura 2 muestra dos pelotones en la calle Carnegie. El primero está detenido por la luz roja en 77-Este y para el segundo acaba de encenderse la verde en 71-Este. Primero los cabezas de este último empiezan a avanzar hacia el cruce de 71-Este; luego, se mueven los automóviles que les siguen. Este movimiento inicial recorre longitudinalmente el pelotón como una onda. Al final, arrancan los últimos automóviles del segundo pelotón.

El pelotón delantero, tan pronto se encienda su luz verde, experimentará también una onda de puesta en marcha que se propagará hasta los vehículos de cola. Si éstos comienzan a moverse antes de que lleguen los cabezas del pelotón trasero, el sistema habrá funcionado perfectamente y los vehículos lo atravesarán sin retrasos indeseables. Supongamos, no obstante, que el pelotón trasero llegue antes de tiempo. Como sus componentes deberán parar, la cola de automóviles se alargará, taponando quizás el cruce con la calle 71-Este durante el siguiente cambio de luces. La situación se convertirá en un verdadero atasco circulatorio.

Yo mismo he quedado varias veces atrapado por un embotellamiento en la calle Carnegie. Un día necesité dos horas para abrirme paso por los seis cruces que hay entre la calle 71-Este y la 86-Este. Una densa nevada había entorpecido la marcha de los primeros pelotones que se adentraron en el tramo. Luego, parte de éstos se vieron detenidos por luces rojas. A causa de la lentitud de la circulación por las calles nevadas, las colas en los cruces se alargaban a medida que se iban juntando otros pelotones antes de que los precedentes hubieran avanzado. La densidad de vehículos no tardó en adquirir unas proporciones colosales: los automóviles se encontraron parachoques con parachoques a lo largo de kilómetros. Como los cruces quedaron obstruidos, el tránsito en aquel triste atardecer de invierno era una materia muy sólida.
urbana se considera un caso de sistema hidrodinámico, he de decirles que el tránsito en aquél triste atardecer de invierno era una materia muy sólida.

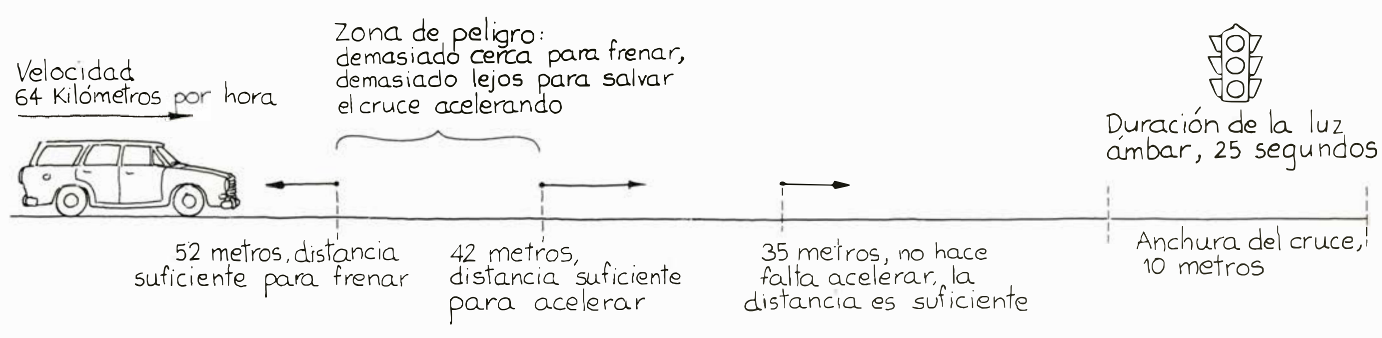
En la situación ilustrada en la figura 2, ¿cuál es la longitud de la porción del pelotón delantero que entorpece al que va detrás? La longitud crítica puede determinarse en función de la distancia entre los cruces y depende de la rapidez con que la onda de puesta en marcha recorre el pelotón de cabeza. Si la onda se propaga a velocidad mucho menor que la velocidad de aproximación del pelotón trasero, hasta un pelotón delantero de escasa longitud puede ocasionar dificultades.

Para medir la velocidad de propagación de la onda de puesta en marcha me estacioné en los cruces de Carnegie durante un período de tránsito lento. Busqué cruces que tuvieran colas de 10 o más vehículos. Cuando el semáforo de una cola se tornaba verde, yo ponía en marcha mi cronómetro; lo paraba cuando comenzaba a moverse el último automóvil de la cola. Además, calculaba a ojo la distancia entre la parte delantera del primer automóvil y la parte posterior del último. Entonces, dividiendo esa distancia por el tiempo transcurrido en mi cronómetro, obtenía un valor aproximado de la celeridad de la onda. Tras un buen número de observaciones encontré que la celeridad media era del orden de cinco metros por segundo; o sea, aproximadamente la mitad de la velocidad de marcha prevista.

Este dato significa que la longitud crítica del pelotón delantero es únicamente una pequeña fracción de la distancia total entre las calles 71-Este y 77-Este. Si este pelotón tiene una longitud mayor que unos tres automóviles, los cabezas del pelotón trasero se verán retrasados. Y, si ambos pelotones son mucho más largos que eso, el pelotón trasero no despejará el cruce de 71-Este antes de que aparezca la luz roja.

Los embotellamientos de la calle Carnegie podrían evitarse si se modificara la secuencia de las luces cuando las colas se alargasen con respecto a la temporización normal de hora punta. Cuanto más nutrido sea el pelotón delantero en 77-Este, menor margen inicial debe darse al pelotón trasero. Y, si la longitud del pelotón delantero es del orden de un tercio de la distancia entre los dos cruces, ambos semáforos deben ponerse verdes a la vez. Con un pelotón delantero más largo todavía, la luz verde de la calle 77-Este ha de encenderse antes que en la 71-Este, orden inverso al de la temporización normal de hora punta. De ese modo, el margen inicial dado al pelotón delantero permitirá que la onda de puesta en marcha alcance a los últimos vehículos antes de la arribada de los cabezas del pelotón posterior.

En la figura 5 aparece una fórmula que da la longitud de los pelotones para la cual hay que cambiar la secuencia normal. La longitud del pelotón (x) está expresada como fracción de la distancia entre los cruces (d) y depende de la velocidad de la onda de puesta en marcha (v_1) y de la velocidad de marcha normal en el tramo (v). Cuando los pelotones sean más cortos que el valor de conmutación, se concederá el margen inicial al pelotón trasero; si más largos, al delantero. Si ambos tienen unas longitudes iguales a la de conmutación, las luces verdes se encenderán a la vez.



4. Alternativas para un conductor que se acerca a una luz ámbar

Longitud de los pelotones para cambiar la secuencia:

$$\frac{x}{d} = \frac{v_1}{(v_1 + v)}$$

Distancia de frenado mínima:

$$d = \frac{v_0^2}{2a}$$

Distancia máxima para atravesar acelerando una luz ámbar:

$$d = v_0 t + \frac{1}{2} a t^2 - s$$

Distancia máxima para atravesar a velocidad constante una luz ámbar:

$$d = v_0 t - s$$

5. Fórmulas para el flujo del tránsito

Hay sistemas de semáforos que pueden cambiarse secuencialmente cuando las colas se alargan en exceso. Ante tal eventualidad, convendría que el sistema dispusiera de una programación preestablecida. En algunos sistemas, la secuencia de las señales está afectada por dispositivos que se depositan en la calzada para captar y registrar la circulación de vehículos. Si la calle Carnegie poseyera un sistema de ese tipo, mis compañeros de cautiverio y yo podríamos habernos ahorrado el sufrimiento que supone recorrer sólo cinco manzanas en dos horas.

Durante las últimas horas de la mañana y primeras de la tarde, la calle Carnegie se divide en tres carriles para cada sentido de circulación. La temporización de las luces ha de ser distinta de la conveniente para la hora punta del atardecer, de suerte que el tránsito que se encamina al oeste no se vea detenido demasiadas veces. Para estudiarla repetí las mediciones de los ciclos de señales aproximadamente a mediodía. Los resultados se exponen en la figura 3, en la que la luz verde en 71-Este se enciende en el instante 10 segundos, tiempo que servirá de referencia para los ciclos de señales en los cruces.

En todos estos cruces, la luz de cada color dura tanto como en la hora punta, pero los intervalos de tiempo entre la aparición de la luz verde en los cruces sucesivos no son los mismos. Por ejemplo, durante la hora punta la luz verde en 79-Este se enciende 15,4 segundos antes que en 82-Este. Cuando el tránsito es bidireccional, el sistema debe sincronizarse para reducir al mínimo el retraso en ambos sentidos de circulación sin castigar al tránsito perpendicular. Al mediodía, el verde en las calles 82-Este y 83-Este está aún encendido cuando se enciende en la 79-Este. Este sistema facilita que el tránsito hacia el oeste fluya ininterrumpidamente a través de los cruces.

En la figura se representan "bandas directas", habituales en estas gráficas de distancia-tiempo. Una banda directa indica de qué modo un pelotón que viaje a velocidad constante cruzará sin interrupción una sucesión de intersecciones. Aunque los automóviles pasen los semáforos a velocidades diversas, las bandas directas sirven a quienes han de decidir cómo sincronizar las luces del tránsito. En la ilustración se ven dos conjuntos de bandas directas, una para la circulación hacia el este (las que ascienden hacia la derecha) y otra para la circulación hacia el oeste (las que descienden hacia la derecha).

Una banda directa se compone de dos líneas paralelas trazadas con la separación que permita la gráfica de una ruta de tránsito urbano. Una de las orientadas hacia el este en la ilustración comienza con la aparición de la luz verde en la calle 71-Este. Una recta pasa por las fases verdes de los otros cruces hasta la 86-Este. (Esta línea es una recta porque representa un tránsito que se mueve a velocidad constante.) Como el tránsito ha de circular por la ruta tan deprisa como pueda, sin rebasar el límite de velocidad, la recta será lo más abrupta posible y evitará, además, todas las fases rojas porque representa una circulación sin detenciones. Así, roza los finales de éstas en las calles 82-Este y 83-Este.

El lado derecho de la banda directa es paralelo al izquierdo porque representa los automóviles a la cola del pelotón, cuyos componentes viajan todos a la misma velocidad. Esta recta se dibuja también evitando todas las fases rojas. En la ilustración, el lado derecho de la banda directa orientada al este roza el comienzo de la señal roja en la calle 86-Este.

Supongamos que el tránsito de mediodía encaminado hacia el este por Carnegie conste de sólo pelotones. La velocidad de marcha de éstos puede calcularse mediante la pendiente de las rectas de las bandas directas del este. Si el tránsito ha de discurrir sin interrupciones, el pelotón arrancará con la aparición del verde en 71-Este y recorrerá los semáforos a una velocidad de 9,6 metros por segundo (equivalente a 34,5 kph).

En el lado derecho de la banda directa se indica cuánto debe tardar la cola del pelotón en atravesar el cruce de la calle 71-Este. Evidentemente, no más de 26 segundos tras el inicio de la fase verde. Si algún vehículo del pelotón lo hiciera más tarde, una luz roja lo detendría en la 86-Este, si no antes. La banda directa para una hora punta es mucho más ancha, abarcando toda la fase verde en 71-Este. Así pues, cuando la red de semáforos de Carnegie cambia a la temporización de hora punta, no se alteran las velocidades de los pelotones; al contrario, éstos pueden ser más largos.

La ilustración permite determinar la distancia entre pelotones contiguos. Comencemos en el lado izquierdo de una banda directa hacia el este, por ejemplo, en el instante de referencia correspondiente a la aparición del verde en la calle 71-Este. En el eje vertical se lee la distancia hacia la izquierda de la banda entre las calles 83-Este y 86-Este para ese mismo instante. La lectura (unos 700 metros) es la distancia entre los automóviles que encabezan dos pelotones sucesivos. La distancia entre los de cola del pelotón delantero y los primeros del que comienza en 71-Este es de unos 500 metros.

En la ilustración se recoge información sobre las bandas directas hacia el oeste. El lado izquierdo de las bandas representa el movimiento de los automóviles que van en cabeza de los pelotones; el derecho, el de los automóviles de cola. Ambas bandas directas, hacia el oeste y hacia el este, tienen aproximadamente tamaños iguales, lo que significa que el sistema de semáforos está concebido para que la circulación transcurra con la misma facilidad en los dos sentidos a esa hora del día.

¿Cuál será una duración idónea de la luz ámbar? Suficientemente corta para no retardar indebidamente el tránsito y suficientemente larga para permitir a cualquier automovilista detenerse o atravesar el cruce antes de que se encienda la luz roja. La duración conveniente dependerá, en parte, del límite de velocidad en cada calle, que determina la capacidad de un conductor para no hallarse en el cruce cuando se encienda la luz roja. Cuanto mayor sea el límite de velocidad, más durará la luz ámbar. Relación ésta que, a veces, se pasa por alto al establecer las fases de un semáforo urbano.

Supongamos que una luz ámbar dure 2,5 segundos, que viene a ser la duración mínima que he podido detectar en Cleveland. Imaginemos además que, cuando se encienda, el conductor, aten-

to, se detiene. La distancia mínima necesaria para parar depende de dos factores: la velocidad inicial del vehículo (v_o) y la aceleración (a) debida al frenazo. (Los físicos llaman aceleración a todo cambio de velocidad; en el habla corriente, la disminución de velocidad de un automóvil se llama deceleración.) La distancia de parada se calcula que es el cuadrado de la velocidad dividido por el doble de la aceleración.

La aceleración de un coche al frenar varía con el tiempo atmosférico, el tipo de pavimento, el estado de las cubiertas y frenos y la destreza del automovilista. Esa aceleración se cifra en un comienzo entre los 0,9 y los 3,1 metros por segundo cada segundo, para alcanzar los 3,7 metros por segundo cada segundo en los últimos momentos de la frenada. En un apuro podría llegar hasta los 6,1 metros por segundo cada segundo, pero frenar con esa intensidad es intranquilizador y desagradable.

Tomemos una aceleración de 3,1 segundo por segundo cada segundo. Supongamos que el conductor marche a 64 kph, que equivale a 17,9 metros por segundo. Entonces, la distancia mínima de parada será de 52 metros y, si el conductor no ha de penetrar en el cruce, no deberá estar más cercano a éste que esa distancia. Si lo está, no tendrá salida más sensata que seguir adelante por el cruce sin tratar de frenar. Acaso llegue a acelerar para no encontrarse en el cruce cuando se encienda la luz roja.

¿Qué posibilidades tendrá el conductor acelerando en tales circunstancias? Lo que consiga dependerá no sólo del poder de aceleración del vehículo, sino también de la distancia que inicialmente lo separe del cruce. Si ésta es excesiva, no podrá llegar al otro lado antes de que se encienda la luz roja. La fórmula de la figura 5 da la distancia máxima al cruce a la que debe encontrarse el automóvil para que el intento tenga éxito. En la fórmula se establece la duración de la luz ámbar (t) y la aceleración (a) y velocidad (v_o) iniciales del vehículo. También se necesita la anchura (s) del cruce, pues el automovilista debe llegar al lado contrario mientras la luz esté aún encendida.

Lo normal es que un automóvil se acelere a razón de un metro a 2,2 metros por segundo cada segundo. Uno deportivo lo hará, posiblemente, a una rapidez doble. Sea un automóvil acelerado a 2,2 metros por segundo cada segundo. Supongamos que su conductor reaccione instantáneamente a la aparición de la luz ámbar. Admitamos que ésta dura 2,5 segundos y que la anchura del cruce sea de 10 metros. Si el vehículo viaja inicialmente a 64 kph, la distancia entre el mismo y el cruce será de unos 42 metros como máximo; si es mayor, se encontrará en el cruce durante parte de la señal roja.

Pero el conductor quizá decida atravesar el cruce sin acelerar por encima del límite de velocidad. Entonces, si mantiene su velocidad inicial de 64 kph, el vehículo no deberá estar separado del cruce más de 35 metros.

He elegido las cifras anteriores con el propósito de llamarles la atención sobre una cosa. Para que el conductor frene con éxito, su vehículo distará del cruce 52 metros por lo menos. Si lo ha de cruzar rápidamente a la aceleración máxima, no debe estar a más de 42 metros. ¿Qué hacer si se encuentra entre esos límites? En principio, no le es favorable ninguna decisión. Esta zona peligrosa entre ambos límites se agranda para velocidades iniciales del automóvil crecientes. Con los mismos datos que antes, pero con una velocidad inicial de 88 kph, la longitud de la zona peligrosa es de 39 metros. Si el conductor prefiere no acelerar, esa longitud es de 46 metros.

En cierta ocasión, circulando por una carretera con un límite de velocidad de 90 kph, me vi en una situación parecida. Me encontré de repente ante una luz ámbar sin espacio para detenerme ni aceleración para atravesar el cruce antes de que se encendiera la luz roja. Sólo me salvé de un choque seguro gracias a un retraso en el sistema de señales: la luz verde del tránsito perpendicular se encendía aproximadamente un segundo después de que se apagara la ámbar.

En la bibliografía de este mes se citan dos artículos en los que sendos profesores de física analizan el dilema de la luz ámbar corta. Este es un tema del que el lector puede ocuparse examinando los semáforos de su barrio. Pero, por favor, téngase presente que los valores de las aceleraciones que yo he utilizado son hipotéticos. La aceleración de frenada es especialmente cuestionable. Depende del rozamiento entre las cubiertas y el suelo de la calle y, si éste se halla cubierto de hielo, nieve, lluvia o cualquier otra cosa que lo reduzca, el coeficiente de rozamiento será mucho menor que lo normal. Entonces, la distancia necesaria para detener el automóvil será mucho mayor. Por tanto, la zona peligrosa se extenderá y crecerá la probabilidad de choque entre un automóvil y un vehículo que circule por una calle perpendicular.

Libros

Ilustración, modelos de computación para biólogos, lógica matemática y etología

Antonio Domínguez Ortíz, Juan Armengol, José F. Prida y Jorge Sabater

ILUSTRACIÓN Y REFORMA EN LA MANCHA. LAS REALES SOCIEDADES ECONÓMICAS DE AMIGOS DEL PAÍS, por José María Barreda Fontes y José Manuel Carretero Zamora. Prólogo de Manuel Espada Burgos. Instituto "Jerónimo Zurita", del Consejo Superior de Investigaciones Científicas (CSIC); Madrid, 1981; 247 páginas. Pasaron los tiempos en los que se identificaba al siglo XVIII, al Siglo Ilustrado, con la frivolidad, la seudociencia, las rizadas y empolvadas pelucas, los abates mundanos y los experimentos de física recreativa para deleite de los asistentes a tertulias aristocráticas, entre un minué de Mozart y unos cuchicheos galantes a la luz de las arañas, que se quiebra en los cristales de Bohemia. Todo eso fue verdad pero estuvo muy lejos de ser toda la verdad.

Fue aquel un siglo duro y contradictorio, en el que se mejoró la productividad de la tierra, en el que una burguesía de negocios supo compaginar el afán de ganancias con la afición a sólidas lecturas y contagió de sus intereses (y también de su sangre plebeya) a una aristocracia cada vez menos segura de sí misma, de su buen derecho, que aplaudía en *Las bodas de Figaro* su propia caricatura y se aturdía con bailes y mascaradas en las vísperas sangrientas de la Revolución. Siglo contradictorio, que extremó la doctrina del absolutismo monárquico a la vez que creaba los mecanismos que iban a expropiarlo. La identificación del Monarca y el Estado había jugado en un primer momento a favor del primero; luego la balanza se inclinaría hacia el segundo. El Estado se despersonaliza y sus representantes, los grandes burócratas, los ministros, se parapetan tras la autoridad real para hacer su propia política, para realizar las reformas que meditan en la administración y en la estructura social. El Estado se va a poner al servicio de la Sociedad, absorberá las atribuciones que competían a los poderes intermedios, a la Iglesia, a los municipios, y se irá bosquejando esa tutela total que en nuestros días conocemos como "Estado Providencia".

Para llevar a cabo este plan, los gobernantes *ilustrados* necesitarían, más que de una sabiduría teórica, de una ciencia volcada hacia la satisfacción de las necesidades concretas de la Nación. No tuvo el siglo de las Luces genios de la talla de Galileo, Leibniz o Newton, pero sí dispuso de una preocupación pedagógica, un afán por transmitir los conocimientos útiles y divulgarlos por medio de la escuela, las academias, los centros de iniciativa particular, la prensa, que en este siglo fue, en su mayor parte, una prensa económica, una prensa orientada hacia la divulgación de los conocimientos útiles.

Todas estas características, propias a Europa entera, se dieron en la España del siglo XVIII. No podemos suscribir la frase de Ortega de que "nos faltó el gran siglo educador". En España advertimos las mismas preocupaciones, quizá con menos fuerza, más atenuadas que en el resto de Europa, porque aquí el motor básico de la transformación, la burguesía de negocios, la burguesía urbana, era débil. Pero los temas son los mismos que en la Francia de Choiseul, en la Inglaterra de los Pitt, en el Nápoles de Vico y Giannone. Una constatación de la necesidad de reformas, una revolución hecha desde arriba, tomando impulso de la Monarquía y sus agentes, pero nutriéndose de las clases que por sus medios, su cultura y su dedicación al bien público pudieron sacar al país de la inercia y efectuar la gran transformación.

Tras la etapa preilustrada, que comienza a fines del siglo XVII y abarca también el reinado de Felipe V, llegó la plena Ilustración con Fernando VI y, sobre todo, con Carlos III. Ni siquiera el desastroso reinado de Carlos IV agotó el movimiento, y hay que reconocer, con Carlos Seco, con Rumeu de Armas, que la etapa de Godoy, por bajo que fuera su nivel moral, mantuvo el impulso adquirido. En 1808 España contaba con un plantel de instituciones y de hombres de ciencia que resultó arrasado por la agresión napoleónica.

Sin duda, el ápice de aquel movimiento hay que situarlo en la segunda mitad del XVIII, y aunque muchos colaboraron en él su figura más dinámica fue la de D. Pedro Rodríguez de Campomanes, fiscal y más tarde gobernador del Consejo de Castilla, máximo órgano de la política interior española. Campomanes, hombre de vasta ilustración y múltiples iniciativas, se dio cuenta de que la renovación de la sociedad hispana, el aumento de la cultura popular, la elevación del nivel de vida de sus pueblos, la recuperación del retraso que llevaba respecto a los países más adelantados era tarea imposible para un Estado absoluto pero embrionario, dotado de escasos medios de acción. Necesitaba la cooperación de las personas cultas y entusiastas; de aquí surgió la idea de las Sociedades Económicas de Amigos del País, sobre un modelo proporcionado ya en 1763-65 por la Sociedad Vascongada, obra de uno de los grupos más característicos de *ilustrados:* los "Caballeritos de Azcoitia".

La iniciativa de Campomanes encontró amplio eco; lo demuestran el centenar de sociedades que han catalogado George y Paula Demerson; unas fueron creadas por excitación directa de las autoridades madrileñas; otras, fruto de iniciativas locales. Sus fines serían variados: la instrucción del pueblo, reducida a los niveles básicos y con vistas a proporcionarle una utilidad inmediata; destierro de la ociosidad mediante la implantación de industrias rurales, sobre todo hilaturas que proporcionaran materia prima a las fábricas de tejidos; mejoras agrícolas, no por transformaciones estructurales sino por introducción de nuevos cultivos, de aparatos

sencillos, nuevos métodos de injerto y poda, cultivo de eriales, etcétera. Fines benéficos también, de asistencia social. En suma, propósitos modestos, concebidos para una España en su mayoría rural y con bajo nivel de vida y de cultura.

José Maria Barreda y Juan Manuel Carretero Zamora han estudiado las Sociedades Económicas fundadas en La Mancha con aquel movimiento. La Mancha era, sin duda, una de las zonas más deprimidas de España. Los manchegos "coronados de doradas espigas" de Cervantes no pasaba de ser una bella frase, un tropo literario. La Mancha era abundante en trigo y viñedo, pero con agricultura extensiva y mala repartición de la propiedad; el latifundio y el arriendo a corto plazo causaban los mismos males que en Andalucía y Extremadura; la miseria perenne de la masa de jornaleros se acrecentaba en los años estériles con la ruina de los pequeños y aun de los medianos agricultores, y entonces se producían las escenas dolorosas que presenció Pons: los hambrientos, arrastrándose en triste éxodo por los caminos para buscar socorro en las limosnas de Toledo y de la Corte. La escasez de una industria rural que sirviera de ocupación alternativa a los campesinos agravaba la dureza de todas esas crisis.

Cinco Sociedades fueron creadas en La Mancha: Toledo, Cuenca, Yepes, San Clemente y Yébenes. Los autores detallan la composición de las mismas; hay un predominio absoluto de miembros del clero, de la nobleza y del funcionariado, con alguna participación de propietarios y rentistas, que en gran parte coincidían con las categorías anteriores. No era de esperar de ellos medidas radicales, más bien una modalidad de las funciones de mecenazgo y beneficiencia que tradicionalmente tenían encomendadas, y aun en este aspecto pecaron más de tacañería que de generosidad. Está claro en la actuación de muchas de aquellas sociedades que la intención de sus miembros era que la masa necesitada dejara de agobiarla con peticiones. En realidad, la única iniciativa de gran envergadura partió del arzobispo Lorenzana: la construcción del hospital de Ciudad Real, lo que fue posible gracias a las cuantiosísimas rentas de la mitra toledana.

A lo largo de las páginas de Barreda y Carretero se advierte el rápido declive que sigue al fugaz entusiasmo inicial. Las sociedades de Alcázar, La Solana y Quintanar de la Orden no llegaron a consolidarse. Las demás llevaron una vida lánguida, reducida a escasas iniciativas. Cuando se indagan las causas de esta decadencia las respuestas son concordantes: falta de recursos, falta de interés, desconfianza hacia su labor. Falta de recursos porque no reciben subvenciones regulares y están reducidas a las cuotas, irregulares y poco cuantiosas, de sus socios; falta de interés en la mayoría de los asociados, que han dado su nombre por compromisos o presiones pero no tienen tiempo ni interés para dedicarles una atención sostenida. Y, sobre todo tras los sucesos revolucionarios de la vecina nación, a partir de 1789 una desconfianza hacia todo lo que fuera iniciativa particular, hacia todo intento de romper o modificar los viejos moldes.

A pesar de todo, las Sociedades Económicas de Amigos del País constituyeron una de las iniciativas más felices de los gobernantes de la Ilustración. Tras el eclipse producido por la Guerra de la Independencia y sus consecuencias inmediatas, renacieron en no pocas ciudades y fueron un fermento de actividad intelectual y cultural dentro del chato y gris panorama provinciano de nuestro siglo XIX. (A. D. O.)

BASIC MICROCOMPUTER MODELS IN BIOLOGY, por James D. Spain. Addison-Wesley Publishing Company; Londres, 1982; 354 páginas. En una época en que la utilización de los ordenadores es poco menos que indispensable para el desarrollo de la mayoría de las diciplinas científicas y el microordenador se está convirtiendo en un elemento de uso personal, empiezan a aparecer libros, como el que se comenta, que introducen su aplicación en determinadas áreas de la ciencia.

Se trata en nuestro caso, fundamentalmente, de un manual. El autor pretende que el lector se familiarice con la aplicación de modelos biológicos mediante el uso de microordenadores, sin detenerse en el marco en que se han desarrollo o en las críticas y limitaciones que pueda tener su aplicación.

El libro consta de tres partes. En la primera (desarrollada en cinco capítulos), se exponen una serie de modelos muy simples que sirven para acostumbrarnos al uso de expresiones matemáticas, transformaciones de los datos y solución numérica de ecuaciones. La segunda parte (14 capítulos) constituye el núcleo central de la obra y en ella se pasa revista a 79 modelos deterministas y multicompartimentales. El agrupamiento de los modelos en capítulos según áreas específicas de su aplicación a la biología permite un tratamiento unitario, a la vez que facilita su empleo por personas interesadas en aspectos concretos. Consta la tercera parte de cuatro capítulos. La más interesante, ofrece modelos probabilísticos con técnicas que se ajustan mucho más a la visión actual de la descripción de procesos biológicos: modelo de Monte-Carlo, cadenas de Markov, teoría de colas y otros.

Se aprecia, en su redacción, un esquema básico, que consiste en exponer brevemente un modelo y proponer a continuación uno o varios ejercicios que representan alguna aplicación del mismo o aspectos útiles desde el punto de vista de su programación. Al final de cada capítulo se encuentran unas conclusiones generales donde se resumen los objetivos perseguidos.

Alguien preguntará por la ausencia de los programas de aplicación al microordenador de los modelos propuestos. El autor considera, creo que acertadamente, que los estudiantes se sienten mucho más estimulados por la posibilidad de desarrollar sus propios programas. En todo caso, se anuncia la existencia de un paquete de programas y ejercicios en disco o cassette aplicables directamente a algunos de los microordenadores de mayor difusión comercial.

Además del objetivo principal de la obra, que se centra en los modelos biológicos *per se* y su aplicación mediante microordenadores, destacaría otros dos aspectos, de evidente provecho para los estudiantes. Por un lado, el libro constituye una valiosa introducción a la aplicación práctica de ciertas nociones elementales de matemáticas. En ocasiones se critica el distanciamiento existente entre la formación matemática que recibe el estudiante durante los primeros cursos de la licenciatura y su viabilidad en el campo de la biología. En segundo lugar, está la familiarización en el manejo de microordenadores y por extensión de aparatos de mayor capacidad que se puede obtener a través de la aplicación de los modelos que aquí se desarrollan. Al igual que con las matemáticas, este aprendizaje se produce a través de la discusión de temas por los que el estudiante se siente inicialmente mucho más atraído.

La principal crítica que puede formularse a la obra es que se trata de una exposición demasiado formal de la modelización, sin tener en cuenta el significado profundo que cada punto de la misma encierra. Diríase que se olvida que todo modelo, como abstracción

mental de una realidad concreta, es un vehículo de ideas y por tanto encierra un significado que trasciende la mera aplicación de fórmulas y la obtención de resultados; si esto se patentizara mejor, el lector, naturalista en ciernes, se sentiría más atraído. No basta con dominar la mecánica de los modelos. Hay que saber cuándo aplicarlos, modificarlos, ver sus limitaciones. Todo modelo surge como un proceso interactivo entre el observador y la realidad a la que se pretende interrogar lo más hábilmente posible. En este proceso, el modelo se va moldeando hasta convertirse en un instrumento que se ajuste a las necesidades del investigador; a veces importa más la evolución seguida que el propio modelo. (J. A.)

PRINCIPIA MATHEMATICA (hasta *56), por A. N. Whitehead y B. Russell. Ed. Paraninfo; Madrid, 1981. Por primera vez se presenta en castellano una parte considerable, sin duda la más interesante, de la monumental obra de Whitehead y Russell, publicada en Cambridge entre 1910 y 1913 y reeditada con algunas leves modificaciones en 1925-27.

La idea genial de Leibniz de reducir la matemática a un juego de símbolos (*Zeichenspiel*) fue apenas desarrollada hasta finales del siglo XIX, aun cuando no deban olvidarse los trabajos de Lambert y de Gottfried Ploucquet, el maestro de Hegel en Tübingen, durante la segunda mitad del siglo XVIII, ni los de Augustus Morgan y George Boole hacia mediados de la siguiente centuria.

El primer logro verdaderamente importante en la formalización de la matemática no aparece hasta 1879, año en que Gottlob Frege publica en Halle su *Begriffsschrift*, acontecimiento que ha sido en ocasiones calificado como el más importante en la historia de la lógica. Mediante la introducción de los símbolos de cuantificación, Frege construyó un lenguaje con capacidad de expresión suficiente para formalizar todas las proposiciones matemáticas, definiendo para este lenguaje unas reglas de deducción, que por primera vez permitían formalizar por completo las demostraciones [véase “Frege”, por Javier de Lorenzo; INVESTIGACIÓN Y CIENCIA, septiembre de 1979].

Coincidiendo casi exactamente con la aparición del *Begriffsschrift*, las investigaciones de Georg Cantor dieron lugar al nacimiento de la teoría de conjuntos, otro notable avance en el campo de los fundamentos de la matemática. La combinación de las ideas de Frege y Cantor –las últimas desarrolladas poco después por Richard Dedekind– permitían aparentemente derivar gran parte, o quizá toda la matemática conocida en un cálculo que incluyese la lógica y la teoría de conjuntos. Tal posibilidad carecía sin embargo de todo interés, ya que pronto pudo demostrarse que tanto el sistema de Frege como la teoría de Cantor adolecían de un gravísimo mal: su inconsistencia, en virtud de la cual en ambos sistemas era un teorema cualquier proposición.

Durante los últimos años del siglo XIX la aparición de paradojas, que en ocasiones eran auténticas contradicciones, conmovió el mundo de los fundamentos de la matemática. Y aunque quizá se haya exagerado el papel que las paradojas representaron en el desarrollo de la lógica formal, es indudable que estimularon las investigaciones en este campo, en el que alrededor de quince años más tarde surgirían resultados de extraordinaria importancia (teoremas de Löwenheim y Skolem), que iniciarían un período de enorme desarrollo de la naciente ciencia. Entretanto, cinco hechos merecen especial mención: las investigaciones de Ernst Schrödel sobre el álgebra de la lógica, la superación por Giuseppe Peano del simbolismo de Frege, la formulación del programa de Hilbert, el nacimiento de la lógica intuicionista de Brouwer y la publicación de los *Principia Mathematica* de Alfred North Whitehead y Bertrand Russell.

Uno de los objetivos de esta obra ciclópea era el de reducir a conceptos puramente lógicos las nociones matemáticas fundamentales, las de número natural y conjunto, idea que ya en parte había llevado a cabo Frege. En un segundo momento se pretendía derivar una parte considerable de la matemática en un cálculo lógico que estuviese libre de contradicciones, lo que ante todo exigiría limitar la noción de conjunto, de forma que la clase A formada por todos los conjuntos no fuese a su vez un conjunto. De lo contrario, el conjunto de las partes de A estaría incluido en A y su cardinal sería en consecuencia igual o menor que el de A, en manifiesta contradicción con un bien conocido resultado de Cantor. Por otra parte, si A era un conjunto también lo sería $B = \{ X \varepsilon A\colon X \not\varepsilon X\}$ (el conjunto de los conjuntos que no son elementos de sí mismos), y se seguiría la contradicción $B \varepsilon B \longleftrightarrow B \not\varepsilon B$.

El problema fue resuelto clasificando los conjuntos jerárquicamente: indi-

viduos son objetos de tipo cero; conjuntos de elementos de tipo cero son objetos de tipo uno; etcétera. Naturalmente, para las relaciones de jerarquía es algo más compleja, al poderse establecer relaciones entre objetos de tipos distintos. Una consecuencia de la jerarquización de los conjuntos es que para todo conjunto X se verifica $X \notin X$, ya que de lo contrario se tendría $\{X\} \subset X$, con lo que un conjunto estaría incluido en otro de orden inferior, obviamente absurdo. Así pues, la clase de todos los conjuntos no puede ser un conjunto y algunas paradojas desaparecen.

La jerarquización de los conjuntos no estaba sin embargo exenta de problemas, uno de los cuales era la aparición de distintos tipos de proposiciones, que conlleva la necesidad de un axioma que permite reducir el orden de una función proposicional y cuya introducción obedece exclusivamente a razones pragmáticas: evitar las paradojas. En virtud de este axioma, denominado de reducibilidad, "cada función de una variable es equivalente, para todos sus valores, a alguna función predicativa del mismo argumento". Conscientes de la dificultad de la justificación lógica de este axioma, en la introducción a la segunda edición los autores manifiestan haber intentado sin éxito sustituirlo por otro que ofreciese menos objeciones.

La parte de la obra original que incluye esta traducción esta dividida en dos partes. La primera, titulada "Lógica matemática", incluye una sección dedicada a la teoría de la demostración, otra relativa a la teoría de las variables aparentes (ligadas, en la terminología actual) y otras tres en las que se aborda el estudio de las clases y relaciones. La segunda parte, titulada "Prolegómenos al cardinal aritmético", es una introducción a la teoría de los números cardinales y ordinales, desempañando en ella un papel esencial el axioma de reducibilidad.

No parece oportuno presentar aquí objeciones a una obra que ocupa un lugar singular en la historia de la lógica. En el ámbito del programa de Hilbert era importante poner de manifiesto la posibilidad de presentar la matemática en un cálculo completamente formalizado, tal y como lo hacen Whitehead y Russell. Sin embargo, la obra se escribe en un momento en que la lógica aún no ha producido un solo resultado interesante y cuando la terminología era aún extremadamente complicada. Lo que, unido al rudimentario estado en que se encontraba la teoría de la demostración, origina complicaciones verdaderamente enormes. De la teoría de tipos, seguramente la aportación más original de los *Principia*, apenas ha quedado otra cosa que la distinción entre conjuntos y clases, mientras que el principio de reducibilidad ha caído por completo en el olvido.

Se trata, pues, de una obra para especialistas, radicando su interés más bien en las introducciones, comentarios y explicaciones que en los desarrollos formales. Sin embargo, para los lógicos provenientes del área de las matemáticas la lectura de los *Principia Mathematica* es difícil. Señalaremos por último el buen hacer del traductor, que ha realizado en forma convincente una labor difícil. (J. F. P.)

BREVE HISTORIA DE LA ETOLOGÍA, por W. H. Thorpe. Alianza Editorial; Madrid, 1982. Debemos entender la etología como una proyección de la moderna zoología, ecología, psicología y hasta de la antropología cultural que pretende estudiar tanto la conducta animal como la humana, desde el punto de vista biológico, contemplando a todos los animales y al hombre como partes integrantes del mismo "continuum".

Esta disciplina tuvo sus antecedentes en la denominada psicología animal del siglo pasado, fundamentada en la filosofía especulativa, por una parte, y lo que podríamos definir como etología descriptiva, por otra. La aparición de la teoría de la evolución de Charles Darwin obligó a adoptar un nuevo enfoque, centrado en la comparación de las distintas tipologías de la conducta animal entre sí, considerando, especialmente, la ubicación de las distintas especies en la escala filogenética, lo que implicaba al hombre en este contexto valorativo.

Esta moderna orientación originó dos corrientes de pensamiento: la etología clásica o europea, fuertemente anclada en la biología, y la psicología comparada, que incluía un grupo heteróclito de especialistas, con predominio de los psicólogos, cuyos fines, similares a los de sus colegas europeos, los han conducido, no obstante, por senderos bastante distintos, ya que partían del concepto mentalista de la naturaleza y el funcionamiento de la "inteligencia humana".

Ambas corrientes convergen hoy. Las discrepancias se diluyen al instituirse una rigurosa metodología científica para explicar su base empírica, asentada, esencialmente, en las modernas

metodologías de observación cuantificable. Ello nos hace suponer que hemos llegado al inicio de su madurez, donde, en un ambiente de sosiego, será posible establecer la unidad y fijar sus verdaderos límites.

La etología es, en España, una disciplina académica nueva, aceptada con recelo por especialistas de diversos campos del pensamiento científico que la estiman poco estructurada y portadora de connotaciones que podrían ser de equívoca interpretación, especialmente si se la contempla desde una perspectiva antropocéntrica.

Está emergiendo con pujanza en nuestras universidades, despertando un interés creciente entre los estudiantes, tanto de biología como de psicología y antropología, por su carácter explicativo de la génesis de los procesos cognoscitivos y conductuales en un campo interdisciplinario. Se trata de una realidad que es preciso asumir sin reticencias, con la objetividad derivada de un conocimiento de sus planteamientos y fines.

La bibliografía etológica en lengua española es todavía muy escasa y me atrevería a calificarla de reiterativa en textos de carácter global; se trata de traducciones de libros angloamericanos o germanos. El esfuerzo de Alianza Editorial en este campo es muy encomiable al haber asumido esta necesidad, traduciendo varias obras clásicas de los principales etólogos e iniciando la penetración en áreas de esta materia que podemos considerar de especialización interdisciplinaria.

La *Breve historia de la etología*, de W. H. Thorpe, de reciente aparición en las librerías, constituye una prueba fehaciente de este anhelo y creo que es la primera historia de la etología escrita en castellano; no obstante su brevedad, se trata de un trabajo de gran interés y rigor, redactado por una autoridad de reconocido prestigio. El autor, biólogo británico especializado, inicialmente, en la fisiología de los insectos, trabajaba en Cambridge a finales de la década de los años 30 en el problema de las adaptaciones respiratorias de las larvas de los parásitos en sus insectos huéspedes; el estudio de esta cuestión, donde confluían la fisiología, el aprendizaje y el instinto, le llevó, de manera paulatina, a la etología, y, concretamente, al estudio de la conducta de las aves, con las que, como buen inglés, estaba familiarizado desde su juventud.

La Universidad de Cambridge, consciente del interés de esta línea de trabajo, le proporcionó, en 1950, un terreno y una vieja mansión en la localidad de Madingley, conjunto que más tarde se convertiría en el "Sub-Department of Animal Behaviour" de Madingley, adscrito a la Universidad de Cambridge. Institución de incuestionable prestigio internacional que bajo la dirección de Thorpe ha formado a etólogos tan notables como Mattews, Peter Marler, Robert Hinde, Alexander Harcourt y otros.

Thorpe, autor de muchos libros sobre comportamiento animal, alcanzó un notable éxito cuando publicó, en 1956, *Learning and Instinct in Animals*, editado por la Universidad de Harvard; obra que, a pesar del tiempo transcurrido desde su aparición, sigue siendo de lectura obligada por todos cuantos se interesan por la ontogénesis de la conducta.

Es tarea ardua escribir un libro, condensado, sobre historia de alguna disciplina científica que sea suficientemente explicativo y que no derive, en demasía, hacia subjetivismos consustanciales con las tendencias científicas del autor; creo que Thorpe lo ha logrado en esta pequeña obra, que conduce al lector por el sendero del desarrollo de esta ciencia desde sus inicios hasta la actualidad.

El libro se divide en tres partes. La primera se centra en los orígenes de la etología; en la segunda aborda la aparición y desarrollo de esta disciplina, y, finalmente, la tercera contempla el futuro de la misma. Tras ahondar en sus inicios remotos, expone cómo en el Renacimiento se establecieron los fundamentos de las futuras ciencias que conocemos con los nombres de ecología y etología. Destaca el impacto de las obras de Lamarck en la biología evolucionista, y, en este contexto, lo que este naturalista francés denominaba "impulso interno" como motor conductual modificador o "transformista" del soma. Buen anglosajón, Thorpe pondera el pensamiento y logros pseudoexperimentales de Spalding en el campo de los "aprendizajes latentes" ya intuidos por Romanes y Morgan dentro del marco de la revolución darwinista, vertiente conductual expuesta por el propio Darwin en 1872 en su *Expresión de las emociones en el hombre y en los animales*.

La escuela etológica americana de principios de nuestro siglo supone, para el autor –criterio que compartimos–, un gran adelanto en los planteamientos que hasta entonces habían condicionado esta ciencia y que podemos resumir en el paradigma "nature versus nurture" con todas sus connotaciones y expectativas referentes a la búsqueda de modelos explicativos de la conducta instintiva, para adentrarnos en un enfoque novedoso que podríamos definir de ponderación del comportamiento animal intacto, tanto en su actuación individual como social, considerar el interés de la primatología como disciplina singular dentro del campo de las ciencias de la conducta, y, finalmente, la imperiosa necesidad de crear una metodología observacional cuantificable que elevara estos balbuceos hasta el nivel de verdadera ciencia. Estos objetivos los debemos a Lashley, Yerkes, Nissen, Bingham y otros.

Rinde homenaje a los etólogos europeos continentales clásicos: Uexkhül, Heinroth, von Frisch, Lorenz, Tinbergen, quienes, moviéndose, generalmentes, en el campo de la psicología gestáltica de Köhler, y, a mi entender, demasiado ensimismados en los complejos paradigmas de la conducta instintiva, han tenido el papel más relevante en el surgimiento de la etología como ciencia, sin olvidar la incidencia decisiva de la neurofisiología que, de la mano de Holst, Hinde y otros, encauzaron lo que Beth, Loeb y Pavlov habían iniciado con la centuria.

Estimo que el autor, muy inmerso en la etología clásica, quizás ha subvalorado el papel, a nuestro entender decisivo, que a esta ciencia han aportado Makkink, Kortlandt y otros con sus etogramas. Estos han creado las bases científicas para una etología observacional, cuantificable, que posteriormente Hinde, Schaller, Marler, Altmann, Sackett y otros muchos supieron asumir con plenitud iniciando un camino que encuentra sus hitos en la ecotología, la etología humana y en una posible sociobiología liberada de sus connotaciones de darwinismo social o del excesivo reduccionismo del norteamericano Wilson.

Se trata, en resumen, de un trabajo muy útil que contempla la etología, en su larga andadura, como una ciencia integradora y esencial para un equilibrado desarrollo de todas las disciplinas que intentan comprender, al animal y al hombre, desde una perspectiva unitaria e integradora.

Y como colofón, desde estas columnas, me permito emplazar a los etólogos españoles para que reflexionen seriamente sobre nuestra dependencia intelectual del extranjero e inicien sus propias publicaciones, producto de sus investigaciones, que nos consta son cada día más numerosas. (J. S. P.)

Bibliografía

Los lectores interesados en una mayor profundización de los temas expuestos pueden consultar los trabajos siguientes:

LA LEY DEL MAR

Ocean Yearbook. Dirigido por Elisabeth Mann Borgese y Norton Ginzburg. International Ocean Institute, University of Chicago Press; Chicago, 1979-82.

Ocean Management: A New Perspective. John M. Armstrong y Peter C. Ryner. Ann Arbor Science Publishers/The Butterworth Group, 1981.

MICROPROGRAMACION

An Overview of Firmware Engineering. Scott Davidson y Bruce Shriver en *Computer*, volumen 11, n.° 5, págs. 21-33; mayo, 1978.

Microprogrammed Control and Reliable Design of Small Computers. George D. Kraft y Wing N. Toy. Prentice-Hall, Inc., 1981.

More Hardware Means Less Software. Robert Bernhard en *IEEE Spectrum*, vol. 15, n.° 12, págs. 30-37; diciembre, 1981.

LA ESCRITURA DE LA CIVILIZACION DEL VALLE DEL INDO

A Dravidian Etymological Dictionary. T. Burrow y M. B. Emeneau. Oxford University Press, 1961.

The Indus Civilization. Sir Mortimer Wheeler. Cambridge University Press, 1968.

The Roots of Ancient India. Walter A. Fairservis, Jr. The University of Chicago Press, 1975.

The Indus Scripts —Texts, Concordance and Tables. I. Mahadevan. Memoirs of the Archaeological Survey of India n.° 77, Manager of Publications, 1977.

A Sketch of Comparative Dravidian Morphology. Kamil V. Zvelebil. Mouton Publishers, La Haya, 1978.

ADN MITOCONDRIAL

Sequence of Introns and Flanking Exons in Wild-Type and *box 3* Mutants of Cytochrome b Reveals an Interlaced Splicing Protein Coded by an Intron. Jaga Lazowska, Claude Jacq y Piotr P. Slonimski en *Cell*, vol. 22, n.° 2, 2.ª parte, páginas 333-348; noviembre, 1980.

Mitochondrial Genes. Dirigido por P. P. Slonimski, P. Borst y G. Attardi. Cold Spring Harbor Monographs, vol. 12, 1981.

Sequence and Organization of Human Mitochondrial Genome. S. Anderson, A. T. Bankier, B. G. Barrell, M. H. L. de Bruijn, A. R. Coulson, J. Drouin, I. C. Eperon, D. P. Nierlich, B. A. Roe, F. Sanger, P. H. Schreier, A. J. H. Smith, R. Staden y I. G. Young en *Nature*, vol. 290, n.° 5806, págs. 457-465; 9 de abril de 1981.

EL FUTURO DEL UNIVERSO

Los Tres Primeros Minutos del Universo. Steven Weinberg. Alianza Editorial. Madrid, 1981.

Time without End: Physics and Biology in an Open Universe. Freeman J. Dyson en *Reviews of Modern Physics*, vol. 51, n.° 3, págs. 447-460; julio, 1979.

Effects of Proton Decay on the Cosmological Future. Duane A. Dicus, John R. Letaw, Doris C. Teplitz y Vigdor L. Teplitz en *The Astrophysical Journal*, vol. 252, n.° 1, págs. 1-9; 1 de enero de 1982.

LA MEMORIA DE LAS AVES QUE ACAPARAN ALIMENTO

Food Storing by Marsh Tits. R. J. Cowie, J. R. Krebs y D. F. Sherry en *Animal Behaviour*, vol. 29, págs. 1252-1259; 1981.

Memory for the Location of Stored Food in Marsh Tits. D. F. Sherry, J. R. Krebs y R. J. Cowie en *Animal Behaviour*, vol. 29, págs. 1260-1266; 1981.

An Experimental Analysis of Cache Recovery in Clark's Nutcracker. S. B. Vander Wall en *Animal Behaviour*, vol. 30, págs. 84-94; 1982.

REACCIONES QUIMICAS OSCILANTES

Biological and Biochemical Oscillators: Proceedings of the Conference on Biological and Biochemical Oscillators, Prague, 1968. Dirigido por B. Chance, E. K. Pye, A. K. Ghosh y B. Hess. Academic Press, 1973.

Chemical Oscillations. G. Nicolis y J. Portnow en *Chemical Reviews*, vol. 73, n.° 4, págs. 365-384; agosto, 1973.

A Systematically Designed Homogeneous Oscillating Reaction: The Arsenite-Iodate-Chlorite System. Patrick De Kepper, Irving R. Epstein y Kenneth Kustin en *Journal of the American Chemical Society*, vol. 103, n.° 8, págs. 2133-2134; 22 de abril de 1981.

Minimal Bromate Oscillator: Bromate-bromide-Catalyst. Miklós Orbán, Patrick De Kepper e Irving R. Epstein en *Journal of the American Chemical Society*, vol. 104, n.° 9, págs. 2657-2658; 5 de mayo de 1982.

HALOBACTERIAS

Purple Membrane of Halobacteria: a new Light-Energy Converter. W. Stoeckenius en *Accounts of Chemical Research*, vol. 13, págs. 337-344; 1980.

Genetic Variability in *Halobacterium Halobium*. F. Pfeifer, G. Weidinger y W. Goebel en *Journal of Bacteriology*, vol. 145, n.° 1, páginas 375-381; 1981.

Characteristics of the Heterotrophic Bacterial Populations in Hypersaline Environments of Different Salt Concentrations. F. Rodríguez-Valera, F. Ruiz-Berraquero y A. Ramos-Cormenzana en *Microbial Ecology*, vol. 7, págs. 235-243; 1981.

Unusual Physical Organization of the Halobacterium Genome. C. Sapienza y W. F. Doolittle en *Nature*, vol. 295, págs. 384-389; 1982.

TEMAS METAMAGICOS

A New UCI Lisp Manual. Dirigido por John R. Meehan. Lawrence Erlbaum Associates, Inc., 1979.

A Micro-Manual for Lisp — Not the Whole Truth. John McCarthy en *History of Programming Languages*, dirigido por Richard L. Wexelblat. ACM Monograph Series, Academic Press, 1980.

TALLER Y LABORATORIO

The Stop-Light Dilemma. Howard S. Seifert en *American Journal of Physics*, vol. 30, págs. 216-218; 1962.

To Stop or Not to Stop — Kinematics and the Yellow Light. J. Fred Watts en *The Physics Teacher*, volumen 19, n.° 2, págs. 114-115; febrero, 1981.